

DISEÑAR UNA METODOLOGÍA PARA DETERMINAR LOS NIVELES DE DEGRADACIÓN
MEDIANTE SEÑALES DE VIBRACIÓN EN RODAMIENTOS, USANDO FUSIÓN DE DATOS
Y CONGLOMERADOS, CON OPTIMIZACIÓN DE PARÁMETROS POR ALGORITMOS
GENÉTICOS.

Andrés Felipe Valencia L.

Proyecto presentado como requisito parcial para optar al título de:
Ingeniero Electricista

Director:
PhD(C), M. Sc. Ing. Mauricio Holguín Londoño
Docente Programa de Ingeniería Eléctrica

Universidad Tecnológica de Pereira
Facultad de Ingeniería Eléctrica
Pereira, Risaralda, Colombia
2018

Agradecimientos...

Quisiera agradecer a mis padres por su apoyo durante este largo camino, su palabras de ánimo y cariño incondicional. A mi director de tesis por su increíble paciencia y dedicación.

TABLA DE CONTENIDO

ÍNDICE DE ILUSTRACIONES.....	6
ÍNDICE DE TABLAS.....	7
PARTE I. INTRODUCCIÓN	8
CAPÍTULO 1. PLANTEAMIENTO	8
CAPÍTULO 2. JUSTIFICACIÓN	9
CAPÍTULO 3. OBJETIVO.....	10
3.1. <i>Objetivo general</i>	10
3.2. <i>Objetivos específicos</i>	10
PARTE II. MATERIALES Y METODOS	11
CAPÍTULO 4. ESPACIOS DE REPRESENTACIÓN.....	11
4.1 <i>Tiempo (T)</i>	11
4.2 <i>Frecuencia (F)</i>	11
4.3. <i>Tiempo - Frecuencia (TF)</i>	12
4.3.1. Coeficientes Cepstrales en las Frecuencias de Mel (MFCC)	13
4.3.2 Transformada de Fourier en tiempo corto (STFT)	13
CAPÍTULO 5. ESTIMACIÓN DE CARACTERÍSTICAS.....	15
5.1. <i>Tiempo (T)</i>	15
5.1.1. Media (Vm).....	15
5.1.2. Desviación estándar (S)	16
5.1.3. Pico positivo máximo (PVM)	16
5.1.4. Raíz media cuadrática (RMS).....	16
5.1.5. Factor de forma (FFO)	16
5.1.6. Factor k (FAK)	16
5.1.7. Histogram upper bound (hub).....	17
5.1.8. Histogram lower bound (HLB)	17
5.1.9. Normal probability density function (NORMALPDF)	17
5.1.10. NORMAL NEGATIVE LOG-LIKELIHOOD VALUE (NLL).....	18
5.1.11. ENTROPIA SHANNON (ENT).....	18
5.1.12. RMS DE PICOS (RMP).....	18
5.1.13 Otras características	18
5.2. <i>Frecuencia (F)</i>	19
5.2.1. Frecuencia media (F1)	19
5.2.2. Varianza (F2).....	19

5.2.3. Asimetría (F3)	19
5.2.4. Kurtosis (F4).....	19
5.2.5. frecuencia central (F5).....	19
5.2.6. Desviación estándar (F6)	20
5.2.7. Raíz media cuadrática (F7)	20
5.2.8 Media geométrica (F8)	20
5.2.9. Desviación media absoluta (F9).....	20
5.2.10 Rango intercuartil (F10).....	21
5.2.10 Otras características	21
CAPÍTULO 6. SELECCIÓN DE CARACTERÍSTICAS	21
6.1. MÉTODOS DE SELECCIÓN DE CARACTERÍSTICAS	22
6.1.1. MÉTODOS COMPLETOS.....	22
6.1.2. MÉTODOS HEURÍSTICOS.....	23
6.1.3. MÉTODOS ALEATORIOS.....	23
CAPÍTULO 7. ALGORITMOS GENÉTICOS	23
CAPÍTULO 8. CLASIFICACIÓN	25
8.1. Cadenas de Márkov	25
8.1.1 El número de estados en el modelo N.....	25
8.1.2 El número de símbolos de observación distintos por estado M.....	26
8.1.3. La distribución de probabilidad de transición $A = \{a_{ij}\}$	26
8.1.4. La distribución de probabilidad de observación de símbolos $B = \{b_j(k)\}$	26
8.1.5. La distribución inicial π	26
8.2 LIBRO DE CÓDIGO	26
8.3. CLUSTERING	28
8.4. K-MEANS	28
CAPÍTULO 9. BASE DE DATOS	29
CAPÍTULO 11. EVALUACIÓN Y SELECCIÓN DE CARACTERÍSTICAS	32
10.1 tiempo (T).....	32
10.2 frecuencia.....	33
10.3 Tiempo-Frecuencia.....	34
CAPITULO 12. CLASIFICACIÓN	35
PARTE III. RESULTADOS, ANÁLISIS, CONCLUSIONES	36
CAPÍTULO 13. MARCO EXPERIMENTAL	36
13.1. Pseudocódigo (Diagrama de flujo).....	37
CAPÍTULO 14. RESULTADOS Y DISCUSIÓN.....	41
14.1 Prueba I (Prueba algoritmo genético).....	41

14.2. Prueba II. (<i>Prueba de libro de código</i>).....	41
14.3. Prueba III (<i>Análisis de tiempo de ejecución</i>).....	41
14.4. Prueba IV. (<i>Algoritmo completo</i>)	42
14.5. Prueba V. (<i>Algoritmo completo</i>).	46
14.6. Prueba VI. (<i>Algoritmo completo</i>)	52
CAPÍTULO 16. CONCLUSIONES.....	59
BIBLIOGRAFÍA	60

ÍNDICE DE ILUSTRACIONES.

ILUSTRACIÓN 1. PROCESADO FFT DE UNA ONDA VIBRATORIA COMPLEJA [41].	12
ILUSTRACIÓN 2. IMPLEMENTACIÓN DE MFCC.	13
ILUSTRACIÓN 3. EXPLICACIÓN GRAFICA STFT [34].	14
ILUSTRACIÓN 4. EXTRACCIÓN DE CARACTERÍSTICAS ESTADÍSTICAS	22
ILUSTRACIÓN 5. FUNCIONAMIENTO BÁSICO DE UN ALGORITMO GENÉTICO.	25
ILUSTRACIÓN 6. LIBRO DE CÓDIGO CODEBOOK.	27
ILUSTRACIÓN 7. CLUSTERING EMPLEANDO K-MEANS.	29
ILUSTRACIÓN 8. BANCO DE PRUEBA [32].	30
ILUSTRACIÓN 9. DISTRIBUCIÓN BASE DE DATOS.	32
ILUSTRACIÓN 10. DIAGRAMA DE FLUJO EXTRACCIÓN DE CARACTERÍSTICAS	37
ILUSTRACIÓN 11. DIAGRAMA DE ENTRENAMIENTO Y VALIDACIÓN	38
ILUSTRACIÓN 12. DIAGRAMA DE FLUJO ALGORITMO GENÉTICO	39
ILUSTRACIÓN 13. DIAGRAMA DE FLUJO ALGORITMO COMPLETO.	40
ILUSTRACIÓN 14 EFICIENCIA VS ITERACIÓN PRUEBA IV	45
ILUSTRACIÓN 15. EFICIENCIA VS ITERACIÓN PRUEBA V	51
ILUSTRACIÓN 16. EFICIENCIA VS ITERACIÓN PRUEBA VI	58
ILUSTRACIÓN 17. COMPARACIÓN RESULTADOS	59

ÍNDICE DE TABLAS

<i>Tabla 1. Otras CARÁCTERISTICAS (T)</i>	<i>18</i>
<i>Tabla 2. Características frecuencia.....</i>	<i>21</i>
<i>Tabla 3. Características tiempo</i>	<i>33</i>
<i>Tabla 4. Características frecuencia.....</i>	<i>34</i>
<i>Tabla 5. Características tiempo-Frecuencia.....</i>	<i>35</i>
<i>Tabla 6. PARámetros de la prueba IV.....</i>	<i>42</i>
<i>Tabla 7. Mejores valores por iteracion de la prueba iv.</i>	<i>44</i>
<i>Tabla 8. Parámetros de la prueba v.....</i>	<i>46</i>
<i>Tabla 9. Mejores valores por iteración prueba V.....</i>	<i>51</i>
<i>Tabla 10. Parametros prueba vl</i>	<i>53</i>
<i>Tabla 11.Resultados prueba VI.....</i>	<i>57</i>

PARTE I. INTRODUCCIÓN

CAPÍTULO 1. PLANTEAMIENTO

En los últimos años el diagnóstico de fallas en máquinas rotativas ha ganado importancia en la industria, como respuesta a la necesidad de disminuir las pérdidas de producción. Debido a la complejidad y estructura de las máquinas rotativas, todos los componentes tienen posibilidad de fallar. Por tanto, es trascendente detectar con exactitud la existencia y severidad de fallas en componentes, para lo cual es importante su diagnóstico y la extracción de características mediante observación de vibraciones [1].

Como sucede a menudo, la gran mayoría de fallos en las máquinas conducen al aumento de los niveles de vibración [2], siendo común el uso de sensores que permitan obtener su comportamiento en un periodo de tiempo. Sin embargo, el análisis de la gran cantidad de datos obtenidos, puede representar importantes tiempos de cómputo, procesos iterativos con información redundante y algoritmos extensos [3]. Existen parámetros determinantes para el uso de algoritmos, es así como en el análisis de fallos la confiabilidad es un factor taxativo el cual ha evidenciado estar estrictamente ligado a la cantidad de datos de entrada [14]. Por lo tanto, de las extensas bases de datos, surge la necesidad de emplear técnicas que faciliten su reducción, pero que conserven la mayor cantidad de información posible.

La transformada fraccional de Fourier (FRFT) ha sido reinventada y usada en distintas áreas, tales como reconocimiento de patrones [4, 5], procesamiento de imágenes [6], estimación de parámetros [7] y filtrado [8]. La FRFT es una transformación tiempo-frecuencia, gracias a su espacio de representación es adecuada para ser usada con señales no periódicas, además, es más flexible que la transformada tradicional de Fourier y la transformada wavelet, debido a la libertad de ajuste del parámetro de rotación angular (p) [9], sin embargo, debido a su naturaleza heurística es necesario repetir el entrenamiento de máquina hasta obtener un valor p óptimo. Sumado a esto, la dimensión del vector de características pre procesado por la FRFT debe ser reducido para disminuir la cantidad de datos de salida [9].

En el diagnóstico de fallas, se ha encontrado que las características de las “fallas raíz” tienden a mejorar la exactitud [1] y resulta de gran utilidad realizar una caracterización seguida de un proceso de selección, para encontrar dichas particularidades. Una de las técnicas que facilitan la realización

de esta tarea son los algoritmos genéticos, los cuales tienen ventaja sobre la selección y construcción de características, sin embargo, en el proceso de extracción, el resultado de estos algoritmos converge a una solución mediante el establecimiento de diferentes parámetros de funcionamiento, lo cual puede no garantizar una solución óptima. Esto significa que el resultado del diagnóstico no es confiable debido a que en cada proceso de entrenamiento la solución puede diferir ampliamente [3].

Con este proyecto se busca dar respuesta a la pregunta de investigación de si es posible mejorar la confiabilidad de los métodos para el diagnóstico de fallos, a través del uso de técnicas de extracción y reducción de características, siendo la fusión la principal herramienta usada para este objetivo.

CAPÍTULO 2. JUSTIFICACIÓN

El mantenimiento ha evolucionado de correctivo a preventivo. El mantenimiento correctivo se refiere a realizar un mantenimiento para corregir una falla existente, mientras el preventivo, se refiere a ejecutar un mantenimiento proactivo con el fin de prevenir problemas en el sistema. Por tanto, es de suma importancia el uso de un método de análisis que pueda ser realizado en tiempo real y sin detener el funcionamiento de la máquina [2]. Una de las metas del mantenimiento, es poder pasar de acciones correctivas o preventivas a métodos totalmente predictivos, los cuales permitan detectar fallos en etapas incipientes. En este orden de ideas, es indispensable evitar la evolución de fallos incipientes hacia problemas de mayor envergadura durante la operación, los cuales puedan ocasionar paros de emergencia y tiempos muertos que se vean reflejados en impactos financieros negativos.

Poder conocer el estado de una máquina es determinante en la continuidad de un proceso industrial, por tanto, es indispensable mantener los equipos en buen estado y poder garantizar su funcionamiento durante intervalos específicos de tiempo, estableciendo condiciones de funcionamiento a mediano y largo plazo. La complejidad empresarial actual y el desarrollo tecnológico involucrado, hacen que el mantenimiento se deba estudiar y aplicar con mayor contenido científico, rigurosidad analítica y profundidad [13].

En la teoría de fusión, la información recogida de diferentes sensores puede ser fusionada para obtener una estimación más robusta. Igualmente se pueden combinar y/o seleccionar diccionarios, o espacios de presentación, con el fin de lograr mejoras en desempeño global del diagnóstico. Diferentes estimadores han sido fusionados con el fin de mejorar sus características y lograr evaluar con mayor precisión los estados del sistema [15]. En [16] se presenta la fusión como un método de compresión capaz de ejecutar una clara reproducción de la señal con una cantidad de datos inferior a la original, dando como resultado la mejora del rendimiento de los algoritmos usados en el análisis de vibraciones.

Usualmente cuando se habla de caracterizar una base de datos existen herramientas que son de gran utilidad para lograr la reducción del tamaño de las mismas, esto con el fin de disminuir tiempos de cómputo y robustez de los algoritmos, la discretización es de gran utilidad cuando se desea trabajar con extensas cantidades de datos, sin embargo, cuando se discretiza una señal puede perderse información valiosa, por eso debe contarse con algoritmos de apoyo que faciliten la reducción y conserven la mayor cantidad de información posible [18]. Como respuesta a la necesidad de disminuir la pérdida de información en el proceso de discretización, se plantea el uso de fuzzy clustering como método de conglomerados [19].

Por tanto, con este proyecto de grado se busca desarrollar una metodología a nivel de fusión de datos, empleado algoritmos genéticos, y a nivel de discretización por métodos de conglomerados (clustering) que permitan mejorar la confiabilidad en el diagnóstico de fallos sobre señales de vibración procedentes de rodamientos.

CAPÍTULO 3. OBJETIVO

3.1. Objetivo general

Establecer una metodología, usando técnicas de fusión de datos mediante algoritmos genéticos y a nivel de discretización por métodos de conglomerados, que establezca los estados de salud de máquinas rotativas mediante el estudio de señales de vibración desde sus rodamientos.

3.2. Objetivos específicos

- Indagar las teorías relevantes sobre caracterización, fusión y discretización para la representación de señales de vibración provenientes de diferentes sensores ubicadas en la máquina.
- Desarrollar una metodología con base en teoría de fusión, que permita optimizar la efectividad del proceso de caracterización.
- Optimizar el proceso de discretización y conglomerados para reducir el tiempo de computo, empleando algoritmos genéticos.
- Integrar las etapas anteriores con el entrenamiento.
- Validar los resultados obtenidos.

PARTE II. MATERIALES Y METODOS

CAPÍTULO 4. ESPACIOS DE REPRESENTACIÓN.

Las señales pueden representarse de diversas maneras, considerándolas vectores y utilizando distintas bases de espacios vectoriales. Los espacios de representación están constituidos mediante un proceso de modelamiento de la señal bajo una variable espacial determinada, este procesamiento de señal es de suma importancia dado que determina los parámetros bases para el desarrollo de estudio de características del modelo. Dos de los espacios comúnmente usados para la representación de señales son: la representación temporal usual que usa la base canónica y la representación en la frecuencia, obtenida por medio de la transformada de Fourier. A continuación, se muestran diferentes espacios de representación [23].

4.1 Tiempo (T)

Es la representación más natural de una señal de vibración la cual se compone de una señal x , que puede ser analizada como una suma de funciones seno provenientes del sistema, la cual está compuesta por diferentes frecuencias (f_i), amplitudes (A_i) y fases (φ_i).

$$x(t) = \sum_{i=1}^n A_i \sin(2\pi f_i t + \varphi_i) \quad (2.4.1)$$

El subíndice (i) representa cada elemento que hace parte del sistema. Cuando se pasa del dominio temporal a otro espacio de representación, se hace necesario el empleo de funciones ortogonales que permitan transformar la señal $x(t)$ a otro espacio de representación [26].

4.2 Frecuencia (F)

En la gran mayoría de los casos las señales se encuentran representadas en el dominio del tiempo, dado que estas son capturadas directamente de la máquina. Sin embargo, la representación en el dominio de la frecuencia es ideal para las señales de vibración por su naturaleza misma. Para ello se emplea la gráfica de amplitud frente a frecuencia que es conocida con el nombre de espectro. Esta es la mejor herramienta que se tiene actualmente para el análisis de vibraciones en maquinaria [27].

El matemático francés Jean Baptiste Fourier encontró la forma de representar una señal compleja en el dominio del tiempo por medio de series de curvas sinusoidales con valores de amplitud y frecuencia específicos.[28] Un analizador de espectros que trabaja con la transformada rápida de Fourier captura

una señal de una máquina, calcula todas las series de señales sinusoidales que contiene la señal compleja y por último mostrarlas de forma individual en una gráfica de espectro, como se muestra en la *ilustración 1* [28].

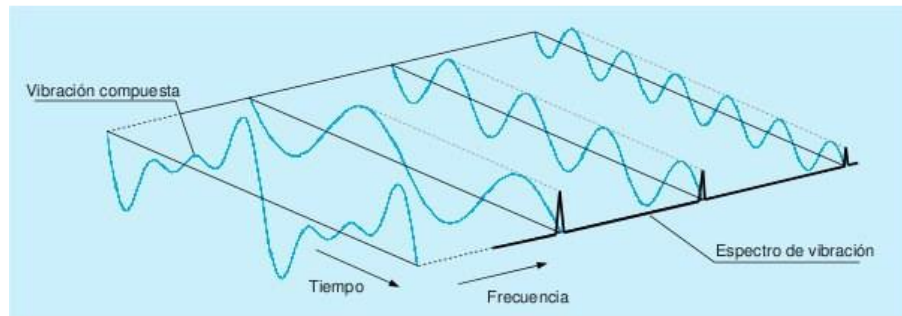


ILUSTRACIÓN 1. PROCESADO FFT DE UNA ONDA VIBRATORIA COMPLEJA [41].

4.3. Tiempo - Frecuencia (TF)

Una representación TF consiste en una superficie tridimensional cuyos ejes son el tiempo y la frecuencia como se observa en *Ilustración 1*, además en la que cada par t-f se dispone de un valor de amplitud al que se le puede llamar energía del punto. Con estos puntos sirve para conocer la forma de la señal en el plano TF y obtener el valor de la concentración de intensidad en cada punto de la distribución global para apreciarla evolución de dicha amplitud tanto en el tiempo como en la frecuencia, aprovechando las características producidas por la concentración de la energía en dos dimensiones (tiempo y frecuencia) en vez de solo una (tiempo o frecuencia). Las transformaciones TF se distribuyen en paramétricas y no paramétricas [30].

- **Paramétricas** son las que utilizan la información de la señal junto con modelos de regresión para obtener el mapa 2D.
- **No paramétricas** son aquellas que realizan una transformación de la señal por tramos, mediante ventanas de selección.

En este trabajo por la segmentación se consideran las transformaciones no paramétricas debido a que requieren un menor tiempo de cómputo que las paramétricas, con la desventaja de que la selección de la ventana adecuada es un parámetro que se debe a justar heurísticamente. Existen diferentes alternativas como solución a la caracterización tiempo-frecuencia con excelentes resultados al

momento de caracterizar señales de vibración como los Coeficientes Cepstrales en las Frecuencias de Mel o MFCCs.

4.3.1. Coeficientes Cepstrales en las Frecuencias de Mel (MFCC)

Los coeficientes cepstrales en las frecuencias de Mel o Mel Frequency Cepstral Coefficients (MFCC) es una representación paramétrica basada en la frecuencia de la percepción auditiva humana. MFCC tiene dos filtros lineales distribuidos para frecuencias inferiores a 1 kHz y logarítmicamente distribuidos para frecuencias superiores a 1kHz [25]. Esta distribución se conoce como la escala de Mel. Los MFCCs se calculan comúnmente de la siguiente manera:

- Separar la señal en pequeños tramos.
- A cada tramo aplicarle la Transformada de Fourier discreta y obtener la potencia espectral de la señal.
- Aplicar el banco de filtros correspondientes a la Escala Mel al espectro obtenido en el paso anterior y sumar las energías en cada uno de ellos.
- Tomar el logaritmo de todas las energías de cada frecuencia mel
- Aplicar la transformada de coseno discreta a estos logaritmos.

La implementación de este proceso se puede resumir tal como se muestra en la *ilustración 2* [25].

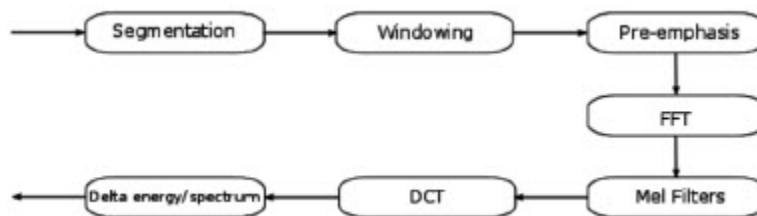


ILUSTRACIÓN 2. IMPLEMENTACIÓN DE MFCC.

4.3.2 Transformada de Fourier en tiempo corto (STFT)

Debido a que el muestreo de la señal con un periodo finito presenta una serie de limitaciones y distorsiones en el análisis mediante la transformada de Fourier. A través de los se ha buscado eliminar la independencia temporal y de este modo disminuir algunas de sus limitaciones. La STFT resuelve el problema del análisis de señales no estacionarias mediante la transformada de Fourier, consiste básicamente en dividir la señal en diferentes partes donde se puede asumir que la señal es estacionaria. Para tal propósito, la señal se multiplica por una función ventana, que inicialmente estará localizada

al inicio de la señal ($t=0$), cuya anchura debe ser igual a la parte de la señal que se quiere considerar estacionaria. [34].

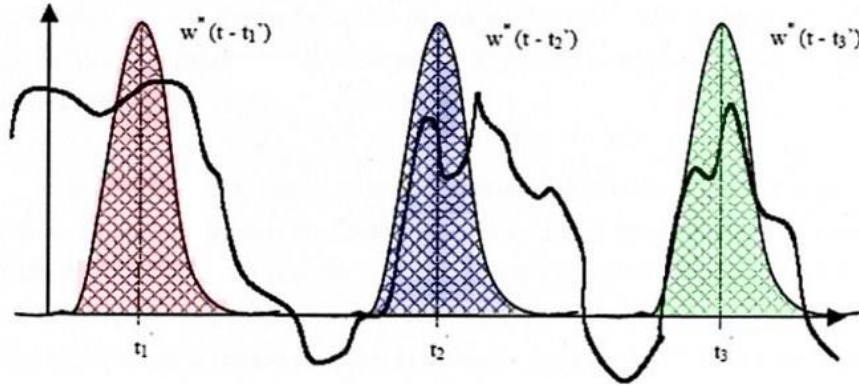


ILUSTRACIÓN 3. EXPLICACIÓN GRAFICA STFT [34].

Asumiendo que la anchura de la ventana es de T segundos, esta función se solapará con la señal para los primeros $T/2$ segundos. La función ventana y la señal se multiplican tomando así solamente los primeros $T/2$ segundos que han sido escogidos de la señal. El resultado de todo esto es una nueva señal producto de la función ventana y de la señal original a la que se le aplica la transformada de Fourier. Como esta hace parte de unas señales estacionaria, los resultados obtenidos mostrarán la representación en frecuencia exacta de los primeros $T/2$ segundos. Posterior a esto consistirá en desplazar esta ventana a una nueva localización a la que de nuevo se le aplicará la transformada de Fourier. De esta forma, se irá moviendo la ventana a lo largo del tiempo hasta que se haya recorrido toda la señal. Una explicación grafica la podemos visualizar en la **ilustración 3**, además que se puede resumir en la siguiente ecuación [34].

$$STFT_{xw}(t',f)=\int [x(t).w^*(t-t')]e^{-j2\pi ft}dt \quad (2.1.1)$$

En donde: $x(t)$. Es la señal original, w^* es la función ventana conjugada, f : frecuencia. Y t : tiempo. Para cada instante de tiempo t' y frecuencia f se calcula un nuevo coeficiente de la transformada de Fourier. Además, hay que hacer constar que la función ventana está caracterizada por ser real, simétrica y de duración finita. La representación de la STFT de la señal en tres dimensiones (tiempo, frecuencia y amplitud) denominada espectrograma, con la que no sólo se conocerían las componentes de frecuencia de la señal, sino también su localización en el tiempo.

La STFT se puede discretizar adoptando como nombre La transformada discreta de Fourier en tiempo corto (STDFT) la cual no es más que la DFT, pero con una ventana de selección. De esta forma, se

obtiene un mejor comportamiento a la hora de llevar a cabo un análisis tiempo-frecuencia de las señales. La ecuación que representa la función dada por $x(n)$ en un tiempo n se define como:

$$X(e^{j\omega k}) = \sum x(m)w(n-m) \quad (2.1.2)$$

Donde

$$\omega k = 2\pi k/N \text{ con } k=0,1,2,\dots,N-1 \quad (2.1.3)$$

ωk es frecuencia en radianes; N es el número de muestras; $w(m)$ es la ventana simétrica seleccionada.

CAPÍTULO 5. ESTIMACIÓN DE CARACTERÍSTICAS.

La etapa de estimación de características consiste en calcular y agrupar las cualidades de las señales de vibración, con el fin de crear una representación manejable y significativa de la señal de vibración original, que permita maximizar el éxito potencial de la fase de clasificación y a su vez el rendimiento global del sistema.

Es importante mencionar la importancia que tiene la fase de estimación de características en el desempeño y rendimiento general de la etapa de clasificación, ya que esta etapa extrae la información útil y comprimiendo los datos que representa una señal sin pérdida de información relevante. A continuación, se muestra el cálculo de las características representadas a partir de su espacio de representación:

5.1. Tiempo (T)

5.1.1. Media (Vm)

Es una medida de tendencia central que en algunos casos puede representar por sí sola todo el conjunto, y se puede calcular como el promedio de un conjunto de valores [35].

$$Vm = \frac{\sum_{n=1}^N x(n)}{N} \quad (2.5.1)$$

5.1.2. Desviación estándar (S)

Es considerado como el segundo momento estadístico de una señal. Sirve para cuantificar el grado de variabilidad y dispersión que presentan las amplitudes con respecto al valor medio [21].

$$S = \sqrt{\frac{\sum_{n=1}^n (x(n) - VM)^2}{N-1}} \quad (2.5.2)$$

5.1.3. Pico positivo máximo (PVM) .

Es el nivel máximo de la señal adquirida, es utilizado como el monitoreo de la variación de los niveles de impulsividad, propiciado por impulsos que alimentan de forma gradual la falla [21].

$$PVM = \max |x(n)| \quad (2.5.3)$$

5.1.4. Raíz media cuadrática (RMS).

El valor *RMS* de una señal de vibración representa un promedio del nivel de vibración en el rodamiento, Permite estimar el grado de severidad de una falla cuando se censa un valor y se compara con datos de referencia establecidos previamente en rodamientos sin falla. Realizando medidas periódicas permite visualizar tendencias que facilitan la evaluación de la condición del rodamiento [21].

$$RMS = \sqrt{\frac{\sum_{n=1}^n (x(n))^2}{N}} \quad (2.5.4)$$

5.1.5. Factor de forma (FFO)

Es la relación entre el valor *RMS* y la media de la magnitud de la señal $x(n)$. Este parámetro permite concentrar el valor *RMS* con el fin de minimizar la imprecisión que se crea en el momento de capturar los datos [21].

$$FFO = \frac{RMS}{\frac{1}{N} \sum_{n=1}^N |x(n)|} \quad (2.5.5)$$

5.1.6. Factor k (FAK)

Es un parámetro muy sensible a los impactos, caracterizado por la relación de proporcionalidad existente entre el factor k (*K*) y los valores *RMS* y el valor pico (*Pv*); es decir, en caso de una

disminución o un aumento en el valor *RMS* o el valor pico, el factor *k* disminuirá o aumentará de igual manera [21].

$$FAK = Pv * RMS \quad (2.5.6)$$

5.1.7. Histogram upper bound (hub)

Proporciona un panorama generalizado de los valores demarcados en el extremo superior de un conjunto de muestras; permite evidenciar más fácilmente ciertas preferencias, tendencias, esquemas de comportamiento y pautas de variación de algún valor asociado, con una falla que puede ser difícil de captar, con otros parámetros [21].

$$HUB = \max(x(n)) - 0.3 * \left(\frac{\max(x(n)) - \min(x(n))}{N-1} \right) \quad (2.5.7)$$

Donde $x(n)$ es la amplitud en el punto de muestreo y N es el número de puntos de muestreo.

5.1.8. Histogram lower bound (HLB)

De la misma manera como ocurre con el límite superior del histograma; este permite evidenciar comportamientos, variabilidades, tendencias y valores atípicos relacionados con una falla, que se encuentran acotados en el extremo inferior de un conjunto de muestras [21].

$$HLB = \min(x(n)) - 0.3 * \left(\frac{\max(x(n)) - \min(x(n))}{N-1} \right) \quad (2.2.8)$$

Donde $x(n)$ es la amplitud en el punto de muestreo y N es el número de puntos de muestreo.

5.1.9. Normal probability density function (NORMALPDF)

La función de densidad de probabilidad de una variable aleatoria continua, calcula la posibilidad de que dicha variable tome determinado valor. La función de densidad de probabilidad (PDF) por sus siglas en inglés es positiva a lo largo de todo su dominio.

$$NORMALPDF = f(x|\mu, s) = \frac{1}{s\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2s^2}} \quad (2.2.9)$$

5.1.10. NORMAL NEGATIVE LOG-LIKELIHOOD VALUE (NLL)

Es utilizado como parámetro de entrada junto con las características anteriormente definidas en este documento [21].

$$NLL = \sum_{n=1}^N \text{Log}(\text{NORMALPDF}) \quad (2.2.10)$$

5.1.11. ENTROPIA SHANNON (ENT)

Es una característica que usualmente es conocida como una medida de la incertidumbre de un proceso. Para un conjunto de eventos con función densidad de probabilidad $P\{x_i, i= 1,2,\dots,N\}$, la entropía de Shannon es definida por la ecuación [21].

$$ENT = - \sum_{n=1}^N P(x(n)) \log P(x(n)) \quad (2.5.11)$$

5.1.12. RMS DE PICOS (RMP)

Raíz media cuadrática de los picos o *RMS of Peaks* (\tilde{x}_{peak}): Es una característica derivada de un análisis de estadísticos calculados a partir de partes de la señal de tiempo en la que se tiene en cuenta la respuesta de impulso de impacto. El valor *RMS* de todas las amplitudes de los máximos locales (picos) en la señal de vibración normalizada se expresa a través de la ecuación [21].

$$RMP = \sqrt{\frac{1}{N} \sum_{n=1}^N x_{npeak}^2} \quad (2.5.12)$$

5.1.13 Otras características

Asimetría	$ASI = \sum_{i=1}^N \frac{(x_i - VM)^3}{s^3} \quad (2.5.13)$
Kurtosis	$KUR = \frac{3s^4}{s^4} \quad (2.5.14)$
Clearence Factor	$FCL = \frac{PVM}{\frac{1}{N} \sum_{n=1}^N \sqrt{ x(n) }} \quad (2.5.15)$
Impulse Factor	$\frac{FIM=PVM}{\frac{1}{N} \sum_{n=1}^N x(n) } \quad (2.5.16)$
T17	$T17 = \frac{1}{N} \sum_{n=1}^N \sqrt{ x(n) }^2 \quad (2.5.17)$

TABLA 1. OTRAS CARÁCTERÍSTICAS (T)

5.2. Frecuencia (F)

5.2.1. Frecuencia media (F1)

La frecuencia media representa el nivel de vibración medio de la señal en términos de la frecuencia. Indica la energía de vibración de la señal en el dominio de la frecuencia [36].

$$F1 = \frac{\sum_{k=1}^K x(K)}{K} \quad (2.5.18)$$

5.2.2. Varianza (F2)

La varianza definida como el segundo momento estadístico, es un indicador de la variabilidad de la amplitud de la vibración en el espectro con respecto a la F1. Esta característica se define como [36]:

$$F2 = \frac{\sum_{k=1}^K (x(K) - F1)^2}{K-1} \quad (2.5.19)$$

5.2.3. Asimetría (F3)

Es el tercer momento estadístico de la distribución de las frecuencias y mide la estabilidad del espectro con medida a la F1 [36].

$$F3 = \frac{\sum_{k=1}^K (x(K) - F1)^3}{K(\sqrt{F2})^3} \quad (2.5.20)$$

5.2.4. Kurtosis (F4)

Es el cuarto momento estadístico de la señal de vibración y es útil para caracterizar fallas en rodamientos y engranajes [36].

$$F4 = \frac{\sum_{k=1}^K (x(K) - F1)^4}{K(F2)^3} \quad (2.5.21)$$

5.2.5. frecuencia central (F5)

La frecuencia central muestra los cambios de posición de las frecuencias principales. Es útil como referente de normalización del rango de frecuencias presentes en la señal [36].

$$F5 = \frac{f_k x(k)}{\sum_{k=1}^K x(k)} \quad (2.5.22)$$

5.2.6. Desviación estándar (F6)

La STDF describe el grado de convergencia del espectro de potencia. [36].

$$F6 = \sqrt{\frac{\sum_{k=1}^K (f_k - F5)^2 x(k)}{K}} \quad (2.5.23)$$

5.2.7. Raíz media cuadrática (F7)

La RMSF, al igual que la F5, muestra los cambios de posición de las frecuencias principales. Es un parámetro que define el nivel de vibración global de la máquina, y es la medida Utilizada con las normas internacionales de vibración. [36].

$$F7 = \sqrt{\frac{\sum_{k=1}^K f_k^2 x(k)}{\sum_{k=1}^K x(k)}} \quad (2.5.24)$$

5.2.8 Media geométrica (F8)

La media geométrica de una cantidad finita de números n, es la raíz n -ésima del producto de todos los números [36].

$$F8 = \sqrt[k]{\prod_{k=1}^K x(k)} \quad (2.5.25)$$

5.2.9. Desviación media absoluta (F9)

La desviación absoluta de un elemento de un conjunto de datos es la diferencia absoluta entre ese elemento y un punto dado. Típicamente, la desviación se calcula desde el valor central, se entienda como algún tipo de media, más a menudo la mediana o, a veces la media del conjunto de datos [36].

$$F9 = \frac{1}{K} \sum_{k=1}^K |x(k) - F5| \quad (2.5.26)$$

5.2.10 Rango intercuartil (F10)

El rango intercuartil es la distancia entre el primer y tercer cuartiles . Es algunas veces llamado la dispersión H y es una medida estable de desembolso [37].

$$F10 = Q_3 - Q_1 \quad (2.5.27)$$

5.2.10 Otras características

MEDIA CUADRATICA	$F11 = \frac{\sum_{k=1}^K f_k^2 x(k)}{\sum_{k=1}^K x(k)} \quad (2.5.28)$
MEDIA ARMONICA	$F12 = \frac{N}{\sum_{k=1}^N \frac{1}{x_k}} \quad (2.5.29)$
$F13 = \frac{\sum_{k=1}^K ((x(k) - \frac{\sum_{k=1}^K x(k)}{K})^3)}{K * \sqrt{\frac{\sum_{k=1}^K (f_k - F5)^2 x(k)}{K}}}$	$F14 = \frac{\sum_{k=1}^K ((x(k) - \frac{\sum_{k=1}^K x(k)}{K})^4)}{K * \sqrt{\frac{\sum_{k=1}^K (f_k - F5)^2 x(k)}{K}}}$
$F15 = \sqrt{\frac{1}{K} * \frac{\sum_{k=1}^K x(k) - \frac{\sum_{k=1}^K x(k)}{K}}{\sqrt{\frac{\sum_{k=1}^K (f_k - F5)^2 x(k)}{K}}} * \frac{1}{K}}$	

TABLA 2. CARÁCTERISTICAS FRECUENCIA.

CAPÍTULO 6. SELECCIÓN DE CARACTERÍSTICAS.

Cuando las características de las señales son obtenidas después de ser calculadas, hay una etapa de procesamiento, ya que para efectos de monitoreo en línea, siempre es deseable reducir la gran cantidad de información contenida en la señal de vibración en línea a un índice único o un pequeño número de características que reflejen el comportamiento general de la señal. Este procedimiento es conocido como selección de características de la señal [38].

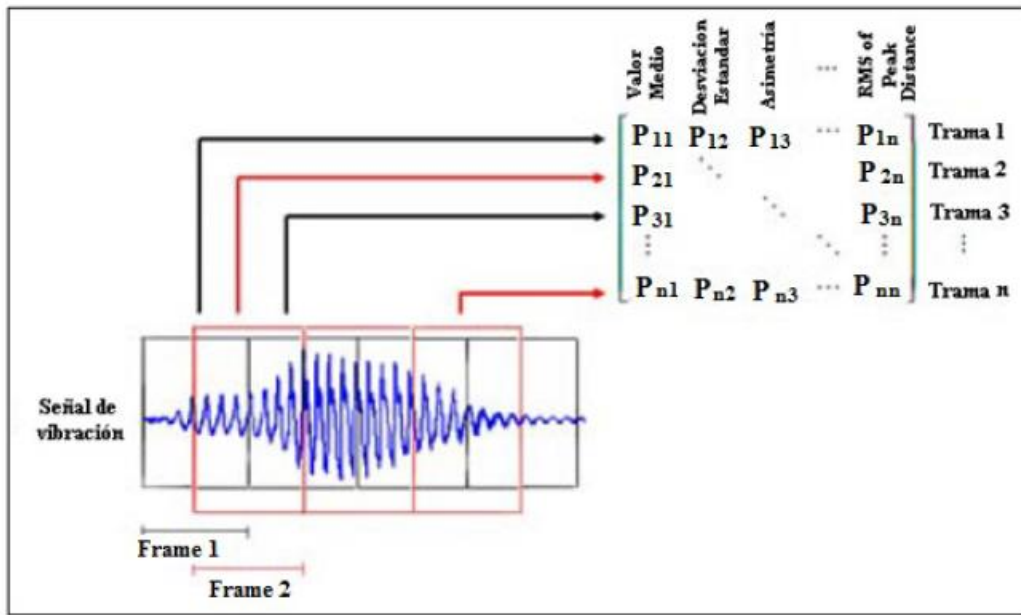


ILUSTRACIÓN 4. EXTRACCIÓN DE CARACTERÍSTICAS ESTADÍSTICAS

El proceso de extracción de características calcula y recolecta en forma de vector compacto, las cualidades de las señales de vibración. La **ilustración 4** muestra gráficamente como es el proceso hecho por el algoritmo de agrupamiento y extracción de datos característicos.

6.1. MÉTODOS DE SELECCIÓN DE CARACTERÍSTICAS.

Los métodos de selección de características se pueden clasificar desde el punto de vista de la manera en que se determina los nuevos subconjuntos a evaluar, lo que conduce a 3 clases métodos conocidos: [39].

6.1.1. MÉTODOS COMPLETOS.

Estos métodos examinan todas las posibles combinaciones de características. Sin embargo, son muy costosos computacionalmente, pero aseguran encontrar el subconjunto óptimo de características. Como ejemplos de estos métodos se pueden encontrar Branch and Bound [40] y Focus [41].

6.1.2. MÉTODOS HEURÍSTICOS.

Utilizan una metodología de búsqueda de forma tal que no es necesario evaluar todos los subconjuntos de características. Ello significa una mayor velocidad del método, ya que el espacio de búsqueda es menor que en los métodos anteriores. Estos métodos no aseguran la obtención del mejor sub - conjunto. A modo de ejemplo es interesante citar en esta categoría los métodos Relief [42] y DTM [43].

6.1.3. MÉTODOS ALEATORIOS

Son aquellos métodos que no tienen una forma específica de definir el subconjunto de características a analizar, sino que utilizan metodologías aleatorias. Con ello se produce una búsqueda probabilística en el espacio de características. El resultado obtenido utilizando este tipo de métodos dependerá del número de intentos, no asegurándose la obtención del óptimo. Pertenecen a este grupo los métodos presentados en LVW [44].

CAPÍTULO 7. ALGORITMOS GENÉTICOS.

Los Algoritmos Genéticos (AGs) son métodos adaptativos que pueden usarse para resolver problemas de búsqueda y optimización. Están basados en el proceso genético de los organismos vivos. A lo largo de las generaciones, las poblaciones evolucionan en la naturaleza de acuerdo con los principios de la selección natural y la supervivencia de los más fuertes, postulados por Darwin (1859).

Estos algoritmos usan una analogía directa con el comportamiento natural. Trabajan con una población de individuos, cada uno de los cuales representa una solución factible a un problema dado. A cada individuo se le asigna un valor o puntuación asociado al fin de dicha solución. En la naturaleza esto equivaldría al grado de efectividad de un organismo para competir por unos determinados recursos. Cuanto mayor sea la adaptación de un individuo al problema, mayor será la probabilidad de que el mismo sea seleccionado para reproducirse, cruzando su material genético con otro individuo seleccionado de igual forma. Este cruce producirá nuevos individuos descendientes de los anteriores, los cuales comparten algunas de las características de sus padres.

De esta manera se produce una nueva población de posibles soluciones, la cual reemplaza a la anterior y verifica la interesante propiedad de que contiene una mayor proporción de buenas características en comparación con la población anterior. Así a lo largo de las generaciones las buenas características

se propagan a través de la población. Favoreciendo el cruce de los individuos mejor adaptados, van siendo exploradas las áreas más prometedoras del espacio de búsqueda. Si el Algoritmo Genético ha sido bien diseñado, la población convergerá hacia una solución óptima del problema [46].

Un algoritmo genético está compuesto por diferentes etapas y se puede aplicar de la siguiente manera:

- Iniciar el algoritmo con una población aleatoria dicha población corresponde a un conjunto de posibles soluciones del sistema.
- Evaluar cada una de las posibles soluciones a través un algoritmo de validación comúnmente conocido como fitness value.
- Seleccionar los mejores “Padres” en base a los resultados obtenidos a través del algoritmo de validación.
- Combinar los mejores “Padres” y obtener un nuevo conjunto de soluciones llamados “Hijos”.
- Aplicar mutación a los “Hijos” con el fin de explorar posibles soluciones que no fueron generadas en la población inicial. El porcentaje de mutación debe ser un porcentaje bajo, esto con el fin de garantizar que los conjuntos óptimos de solución no se vean drásticamente afectados al final de cada iteración. Usualmente el porcentaje de mutación se encuentra en el rango de 1 a 5%.
- Establecer un criterio de parada y evaluar si las condiciones fueron cumplidas.

En la *ilustración 5* se presenta gráficamente el funcionamiento básico de los algoritmos genéticos.

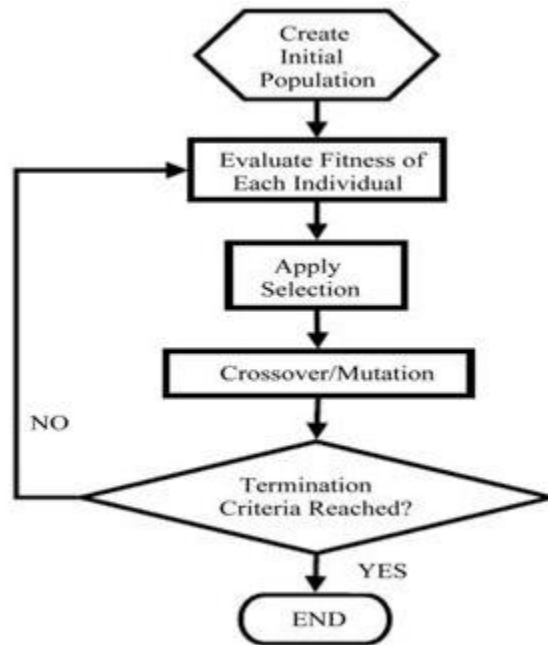


ILUSTRACIÓN 5. FUNCIONAMIENTO BÁSICO DE UN ALGORITMO GENÉTICO.

CAPÍTULO 8. CLASIFICACIÓN.

8.1. Cadenas de Márkov

Las cadenas de Markov son modelos estadísticos de señales observables representadas como un proceso aleatorio paramétrico, es decir, los parámetros pueden determinarse con precisión. Por otro lado, los Modelos Ocultos de Markov (HMM, del inglés Hidden Markov Models) son una extensión de las cadenas de Markov cuando las señales no son observables [21], dando como consecuencia una ruta de estados no es observable, pero la salida, dependiendo del estado, si lo es. Cada estado posee una distribución de probabilidad sobre las posibles señales de salida. Por consiguiente, los modelos ocultos de Markov ofrecen indicaciones sobre el orden de los estados con ayuda de la secuencia de símbolos que estos generan.

Los HMM constan básicamente de 5 elementos para su construcción los cuales son:

8.1.1 El número de estados en el modelo N.

Aunque los estados en los que se encuentra el sistema modelado se consideran ocultos, para muchas aplicaciones prácticas existe alguna significación física asociada a los estados del sistema.

Generalmente los estados están interconectados de tal manera que cualquier estado puede ser alcanzado desde cualquier otro. Todos los estados serán denotados como $S = \{S_1, S_2, \dots, S_N\}$ y el estado en el instante t como q_t .

8.1.2 El número de símbolos de observación distintos por estado M .

Es de tamaño discreto del modelo. Sus símbolos corresponden usualmente a la salida “física” u observable del sistema modelado. Se denotan los símbolos como $V = \{v_1, v_2, \dots, v_m\}$, y su cantidad de símbolos observables como M .

8.1.3. La distribución de probabilidad de transición $A = \{a_{ij}\}$

Usualmente representada por una matriz en donde:

$$A = \{a_{ij}\} = P(q_{t+1} = S_j | q_t = S_i) \quad (2.8.1)$$

8.1.4. La distribución de probabilidad de observación de símbolos $B = \{b_j(k)\}$.

Esta distribución de probabilidad representa la probabilidad de observar el símbolo k estando en el estado j donde:

$$B = \{b_j(k)\} = P(v_k | q_t = S_j), 1 \leq j \leq N, 1 \leq k \leq M \quad (2.8.2)$$

8.1.5. La distribución inicial π Representa el estado inicial del sistema modelado, donde:

$$\pi = \{\pi_i\} = P(q_1 = S_i) \quad 1 \leq i \leq N \quad (2.8.3)$$

Dados valores apropiados para N , M , A , B y π , el HMM puede ser utilizado como el generador de una secuencia de observaciones.

8.2 LIBRO DE CÓDIGO

El libro de códigos nace con la idea de encontrar y cambiar la representación de los vectores característicos, que pueden tomar infinitos valores, por un conjunto finito de vectores que constituyen la señal, donde es comparada con cada patrón en un cierto intervalo, y la señal en cada período es proyectada por un código que indica el patrón que es más parecido a la señal.

Esta cuantificación de vectores trae consigo la idea de clasificar un conjunto de vectores, luego de lo cual se buscarán los mejores representantes para reducir el tamaño de la información a manejar, con el fin de obtener un grupo de vectores representativos llamado libro de código (codebook).

Idea en general trata de clasificar un conjunto de vectores, que en nuestro caso serán las características generadas a partir del espacio de representación, luego de lo cual se buscarán los mejores representantes para reducir el tamaño de la información a manejar. La forma de medir la fidelidad de un cuantificador es determinar el error que éste produce al reemplazar los datos de entrada que recibe por los vectores representantes o codewords, dicho parámetro es llamado error por distorsión. La finalidad de un cuantificador es obtener un conjunto de vectores representativos llamado codebook, que presente el menor error por distorsión [46].

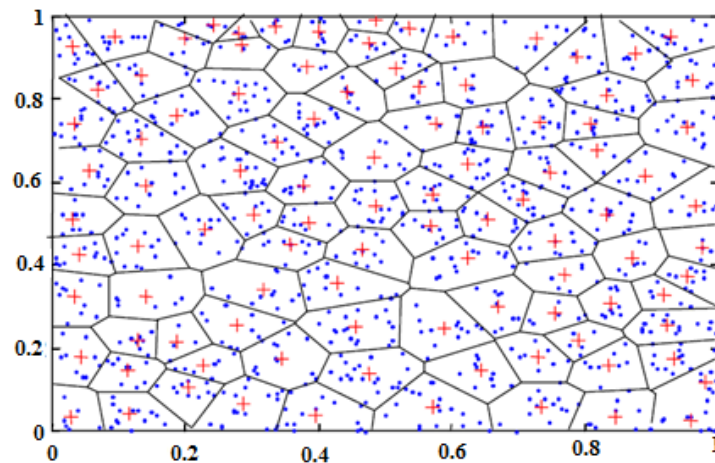


ILUSTRACIÓN 6. LIBRO DE CÓDIGO CODEBOOK.

En general existen dos formas para generar el libro de código:

- **Aprendizaje aleatorio:** Permite a los vectores ser elegidos de manera aleatoria de los datos de entrenamiento y luego son guardados como vectores de códigos.
- **Por clúster:** Agrupan los datos de entrenamiento en grupos mutuamente separados y a cada grupo se le calculan las medias permitiendo minimizar la distorsión promedio, estas medias son guardados como vectores de códigos.

8.3. CLUSTERING

Clustering es el proceso en el cual se agrupan datos en clases, de tal manera que los objetos de un clúster posean un nivel elevado de semejanza entre ellos, y un nivel bajo de semejanza con objetos de otros clusters.

El Análisis Clustering es una técnica descriptiva, a teórica y no inferencial, por lo cual no tiene bases estadísticas sobre las que deducir inferencias estadísticas para una población a partir de una muestra, es un método basado en criterios geométricos y se utiliza fundamentalmente como una técnica exploratoria, descriptiva pero no explicativa [47].

Por otra parte, la solución clúster depende totalmente de las variables utilizadas, la adición o destrucción de variables relevantes puede tener un impacto substancial sobre la solución resultante. Los algoritmos de formación de conglomerados se agrupan en dos categorías:

- **Algoritmos de partición:** Método de dividir el conjunto de observaciones en k conglomerados (clusters), en donde k lo define inicialmente el usuario.
- **Algoritmos jerárquicos:** Método que entrega una jerarquía de divisiones del conjunto de elementos en conglomerados.

Existe una gran variedad de algoritmos de clustering como: K-means, CobWeb, Algoritmo EM (Expectation Maximization), SOM (self-organization maps), entre otros. Nuestro trabajo está enfocado en el estudio del algoritmo de K-means.

8.4. K-MEANS

k-means es un algoritmo de agrupación el está destinado a situaciones en las cuales todas las variables son del tipo cuantitativo, y la distancia cuadrática Euclidea es elegida como medida de diferencia.

$$d(x_i, x_i') = \sum_{j=1}^P (x_{ij} - x_{i'j})^2 = \|x_i - x_i'\|^2 \quad (2.7.3)$$

Se nota que los pesos en la distancia Euclidea pueden ser usados redefiniendo los valores x_{ij} . Los puntos de dispersión pueden ser escritos como:

$$W(C) = \sum_{k=1}^K \sum_{c(i)=k} \sum_{c(i')=k} \|x_i - x_{i'}\|^2 \quad (2.7.4)$$

$$W(C) = \sum_{k=1}^K \sum_{c(i')=k} \|x_i - x_{i'}\|^2 \quad (2.7.5)$$

Donde $\bar{x}_k = (\bar{x}_{1k}, \bar{x}_{2k}, \dots, \bar{x}_{pk})$, es el vector de medias asociado con el k-esimo clúster, y $N_k = \sum I(C_i = K)_{i=1}$. Así, el criterio es asignar las N observaciones a los K clúster de modo que dentro de cada clúster el promedio de las diferencias de cada observación a la media del clúster, definido por los puntos del clúster, sea mínima [48].

Cuando la representación es llevada a cabo utilizando centroides, la inmediatez en la entrega de resultados se convierte en una gran ventaja gracias a su representación de naturaleza gráfica y estadística. Cada clúster por tanto es caracterizado por su centroide, ver **ilustración 7**. que se encuentra en el medio de los elementos que componen el clúster [49].

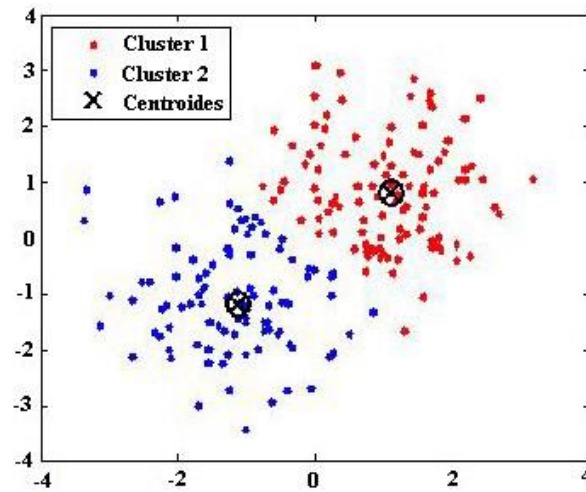


ILUSTRACIÓN 7.CLUSTERING EMPLEANDO K-MEANS.

CAPÍTULO 9. BASE DE DATOS.

La base de datos fue obtenida por Bearing Data Center de la Western Reserve University [32], y esta se encuentra organizada en archivos digitalizados obtenidas de mediciones, con el fin de realizar un

análisis de vibraciones adecuado según el tipo de falla, la cual resulta útil como modelo de referencia a la hora de poder contar con una base de datos y su estudio de cada elemento de una máquina sirviendo de reseñas ante posibles anomalías.

La importancia de contar con estos datos radica en la posibilidad que brinda el poder llevar el monitoreo de su estado general y así poder elaborar la línea de tendencia de los equipos. Con estas lecturas se puede hacer una comparación con datos recientes y de esta manera poder monitorear las frecuencias de elementos críticos como son los rodamientos.

Esta central ha documentado cuidadosamente el estado y el tipo de fallo del rodamiento para cada experimento, utilizando como referencia un motor eléctrico de 2 HP. El motor se conectó a un dinamómetro y a un sensor de par mediante un acoplamiento de auto alineación. Los datos de aceleración se midieron en ubicaciones cercanas y a distancia de los rodamientos. A los rodamientos se les crearon fallas de forma artificial mediante mecanizado por electro-descarga (EDM). Los fallos con un rango de 0,007 pulgadas de diámetro a 0,040 pulgadas de diámetro, con profundidades de 0.011024 pulgadas. Se introdujeron por separado en el anillo interior, los elementos rodantes (es decir, bola) y la pista de rodadura exterior.

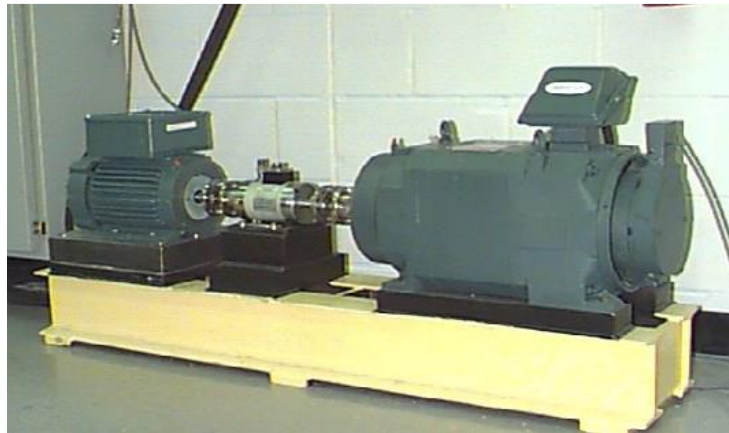


ILUSTRACIÓN 8. BANCO DE PRUEBA [32].

Las especificaciones de los rodamientos empleados en la prueba fueron:

- Diámetro de la bola 0.31260 pulgadas.
- Diámetro de paso 1.537 pulgadas.
- Numero de bolas 9.

- Ángulo de contacto cero (0).

Los rodamientos con falla fueron nuevamente instalados en el motor de prueba y los datos de vibración se registraron para cargas del motor de 0 a 3 HP (velocidad del motor oscilaba entre 1720 a 1797 RPM). La carga se aplicó utilizando el dinamómetro y los datos de velocidad y potencia se recolectaron a través de un sensor. Las señales de vibración se adquirieron con una frecuencia de muestreo de 12 [kHz].

Los registros de las señales de vibración originales están almacenados en archivos con extensión (.mat) agrupados de acuerdo al estado de salud de los rodamientos:

- Fallo en bola (Ball).
- Fallo en la pista externa (Outer).
- Fallo en la pista interna (Inner).
- estado normal (Normal).

El estado de salud a la vez, se subdivide en tres niveles de severidad (N1, N2, N3), cada nivel de severidad (N1, N2, N3) correspondiente a cada tipo de falla (Ball, Outer, Inner y Normal) está compuesto por un conjunto de doce registros correspondientes a cuatro velocidades diferentes, en donde cada registro, almacena un arreglo bidimensional de 40000x1, como se muestra en la *ilustración 9*.

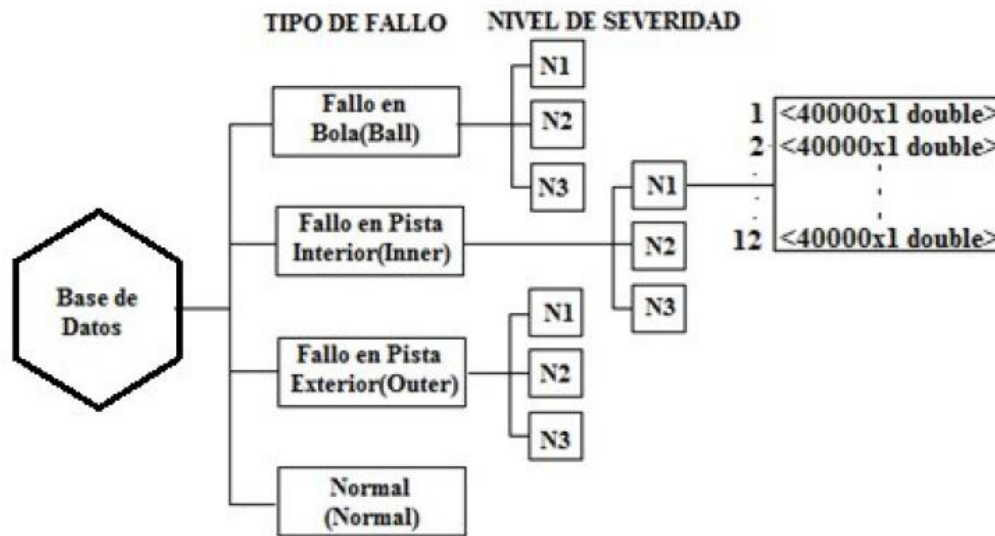


ILUSTRACIÓN 9. DISTRIBUCIÓN BASE DE DATOS.

CAPÍTULO 11. EVALUACIÓN Y SELECCIÓN DE CARACTERÍSTICAS.

Esta etapa consiste en estimar y extraer las características a partir de señales que fueron segmentadas y ventaneadas, para calcular y agrupar las cualidades de las señales de vibración en un formato de vector compacto, con el fin de crear una representación manejable y significativa de la señal de vibración original, que permita maximizar el éxito potencial de la fase de clasificación y a su vez el rendimiento global del sistema.

Luego de la segmentación, sobre cada una de las tramas (frames) de las señales de vibración originales se evaluaron, las siguientes características estadísticas que representan el comportamiento temporal de la señal:

10.1 TIEMPO (T).

En la **tabla 3** se muestran las características de tiempo usadas en el análisis y su respectiva numeración.

1. MEDIA(VM).	9. T17
2. DESVIACIÓN ESTANDAR (S).	10. CLEARANCE FACTOR .
3. ASIMETRÍA	11. IMPULSE FACTOR NORMAL
4. KURTOSIS	12. HISTOGRAM UPPER BOUND (HUB).
5. PICO POSITIVO MÁXIMO (PVM).	13. HISTOGRAM LOWER BOUND (HLB)
6. RAÍZ MEDIA CUADRÁTICA (RMS).	14. NEGATIVE LOG-LIKELIHOOD VALUE (NLL)
7. FACTOR DE FORMA (FFO)	15. ENTROPIA SHANNON (ENT)
8. FACTOR K (FAK)	16. RMS DE PICOS (RMP)
	17. DESVIACION ESTANDAR DEL UMBRAL

TABLA 3. CARACTERÍSTICAS TIEMPO

Para este cálculo no es necesario una transformación previa de su espacio. Por lo cual su representación en el tiempo es la base para los cálculos de las características.

10.2 FRECUENCIA

Las características usadas para este espacio de representación explicadas en el *capítulo 4* las cuales se encuentran ordenadas como se muestra en dicho capítulo. Para este cálculo es necesario una transformación previa de su espacio a través de la transformada rápida de Fourier), Por lo cual su representación en el dominio de la frecuencia es la base para los cálculos de las características, en la *tabla 4* se muestran las características de frecuencia utilizadas y su respectiva numeración.

1. Media	9. Inversa de la raíz cuadrática por la media cuadrática
2. Varianza	10. Desviación estándar sobre la frecuencia media
3. Asimetría	11. F13
4. Kurtosis	12. F14
5. Mediana	13. F15
6. Desviación	14. Media geométrica
7. Raíz Media Cuadrática	15. Desviación Media absoluta
8. Media Cuadrática	16. Rango Inter cuartil

TABLA 4. CARACTERÍSTICAS FRECUENCIA

10.3 Tiempo-Frecuencia.

Para el análisis usado en el espacio tiempo frecuencia se usó la representación paramétrica (MFCC) explicada en el capítulo 4 en la sección (2.4.3.1). El cálculo de las diferentes características usadas para la representación tiempo frecuencia se presenta en la **tabla 5**.

1. $TF1 = \frac{1}{N} * FR_1$	7. $TF7 = Kurtosis(FR_2)$
2. $TF2 = Max FR_1$	8. $TF8 = \frac{\sum FR_2 * FR_2}{2}$

3. TF3=Kurtosis(FR_1)	9. TF9=Kurtosis(WD)
4. TF4=$\frac{\sum FR_1 * FR_1}{2}$	10. TF10=$\frac{1}{2} \sum (WD)^2$
5. TF5=$\frac{1}{N} * FR_2$	11. TF11=$\frac{1}{N} \sum_{n=1}^N WD_n$
6. TF6= MaxFR_2	12. TF12=Max(WD)

TABLA 5. CARACTERÍSTICAS TIEMPO-FRECUENCIA

Donde FR corresponde a cada una de las ventanas (Frames). Y el subíndice corresponde al número de la ventana, por ejemplo FR_n corresponde a la ventana número n.

Donde WD es el espectrograma de las ventanas.

CAPITULO 12. CLASIFICACIÓN

La clasificación del sistema se basa en Modelos Ocultos de Márkov (HMM) los cuales poseen la capacidad de escoger una correcta selección de la arquitectura y de los parámetros del clasificador. Para ello, se cuenta con una base de datos lo suficientemente amplia para el entrenamiento y para la validación del clasificador.

Los datos muestrales dividen en dos partes; la primera parte se utiliza como conjunto de entrenamiento para determinar los parámetros del clasificador con un 70% de los datos y la otra parte, llamado conjunto de validación con un 30%,.

El software para cada fase del calificador está basado en un lenguaje matemático que ofrece un entorno de desarrollo completo con un lenguaje de programación integro como lo ofrece MATLAB, este posee librerías llamadas “toolbox” que son dedicadas al entrenamiento y clasificación de modelos usando Modelos Ocultos de Márkov (HMM), conformados por pequeños scripts y funciones que permiten verificar paso a paso las condiciones del sistema.

Se usa la validación cruzada como herramienta para evaluar los resultados de un análisis obtenido por el entrenamiento a partir de los Modelos Ocultos de Márkov (HMM), garantizando, que son independientes de la partición entre datos de entrenamiento y prueba. La valides en este sistema viene dada por los conceptos de sensibilidad y especificidad de una prueba [27].

PARTE III. RESULTADOS, ANÁLISIS, CONCLUSIONES.

CAPÍTULO 13. MARCO EXPERIMENTAL.

Este trabajo se establece una metodología de análisis y clasificación de fallas mecánicas en rodamientos usando técnicas de fusión de datos mediante algoritmos genéticos y a nivel de discretización por métodos de conglomerados para el análisis de vibraciones haciendo uso de algoritmos genéticos. para lo cual se realiza una caracterización, en tiempo, frecuencia y tiempo-frecuencia de las señales de vibración obtenidas a partir de una prueba del Bearing Data Center de la Western Reserve [32].

El objetivo del análisis de vibraciones es realizar un estudio riguroso de las señales vibratorias con el fin de determinar el estado de salud del equipo y detectar posibles anomalías mecánicas o funcionales de la misma.

Las vibraciones pueden ser entendidas como el movimiento de una partícula o de un cuerpo, o sistema de cuerpos conectados, desplazados desde una posición de equilibrio. Este movimiento se repite después de cierto intervalo, dicho intervalo se conoce como periodo de vibración. La mayoría de las vibraciones son indeseables ya que estas producen pérdidas de energía, desgaste extra, aumento de carga, entre otras [21].

Actualmente los planes de mantenimiento predictivo de los sistemas mecánicos se basan en el estudio de las vibraciones de dicho sistema, utilizando sensores que permiten detectar formas de vibración anómalas en la máquina, dichas formas de ondas son analizadas a través de métodos desarrollados en el análisis de fallas. El análisis de falla es un examen sistemático de la pieza dañada para determinar la causa raíz de la falla y usar esta información para mejorar la confiabilidad del equipo, el análisis de falla está diseñado para identificar los modos de falla (la forma de fallar del producto o pieza), identificar el mecanismo de falla (el fenómeno físico involucrado en la falla), determinar la causa raíz (el diseño, defecto, o cargas que llevaron a la falla) y recomendar métodos de prevención de la falla [22].

13.1. Pseudocódigo (Diagrama de flujo).

- Caracterización de la señal

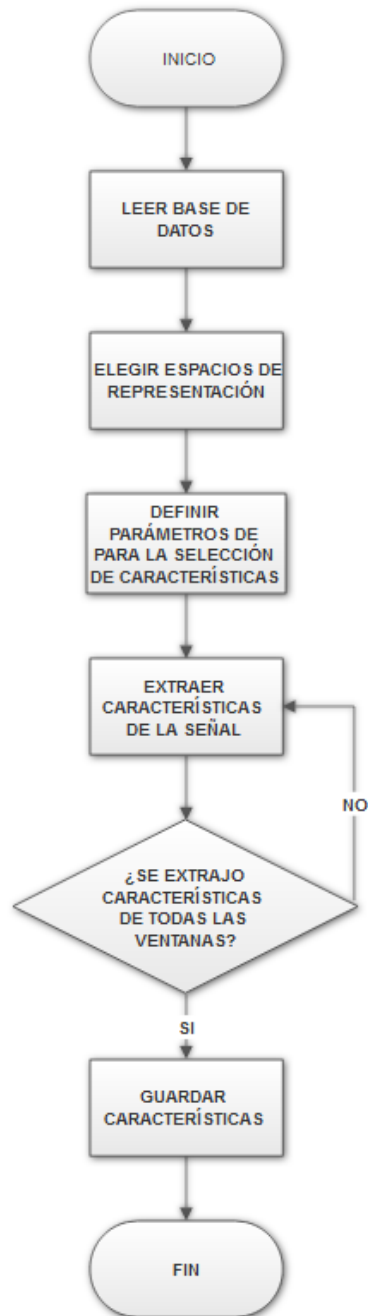


ILUSTRACIÓN 10. DIAGRAMA DE FLUJO EXTRACCIÓN DE CARACTERISTICAS

- **Entrenamiento y validación.**

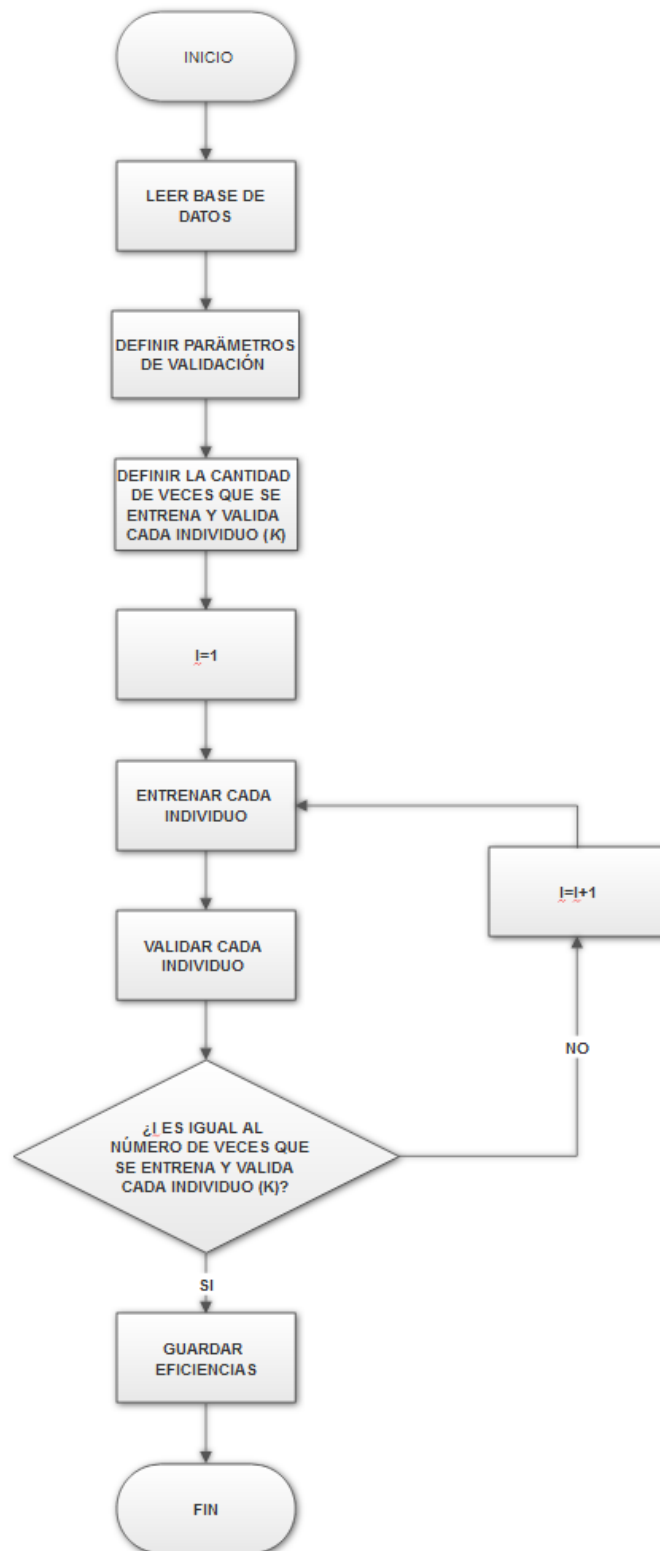


ILUSTRACIÓN 11. DIAGRAMA DE ENTRENAMIENTO Y VALIDACIÓN

- **Algoritmo Genético.**

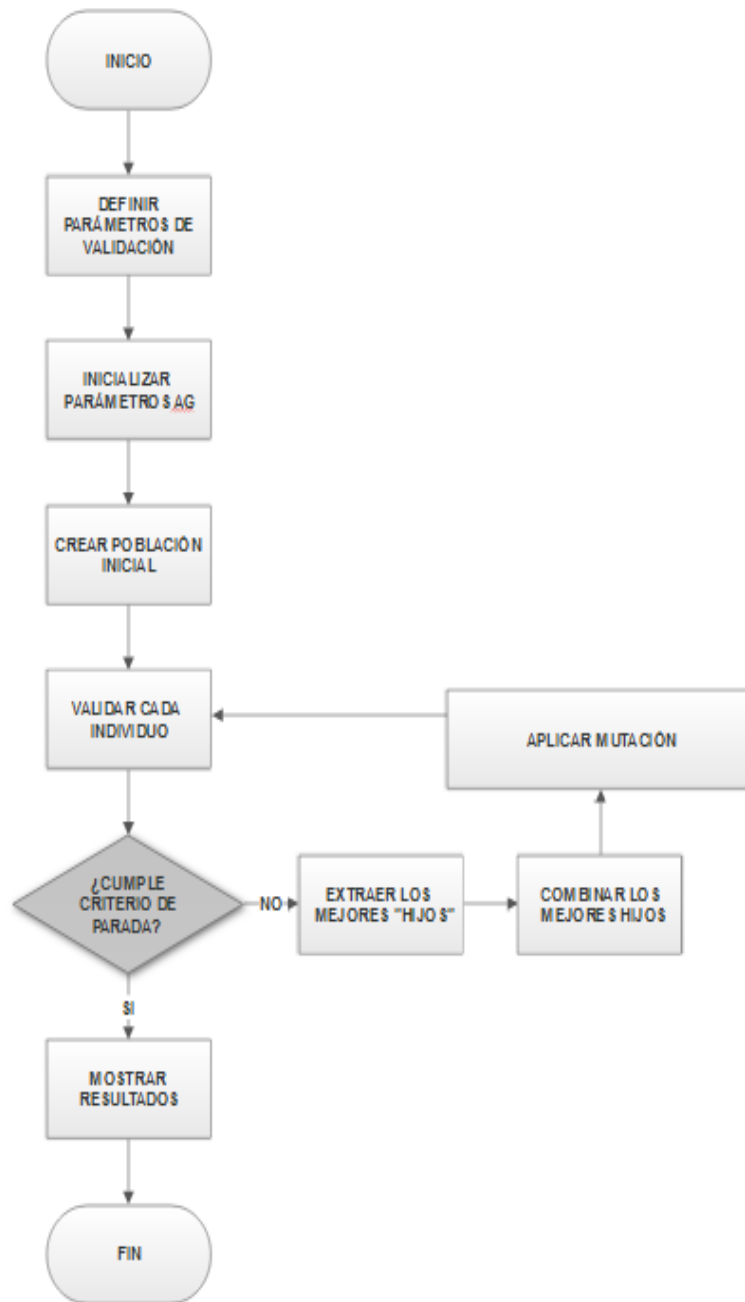


ILUSTRACIÓN 12. DIAGRAMA DE FLUJO ALGORITMO GENÉTICO

- Algoritmo completo.

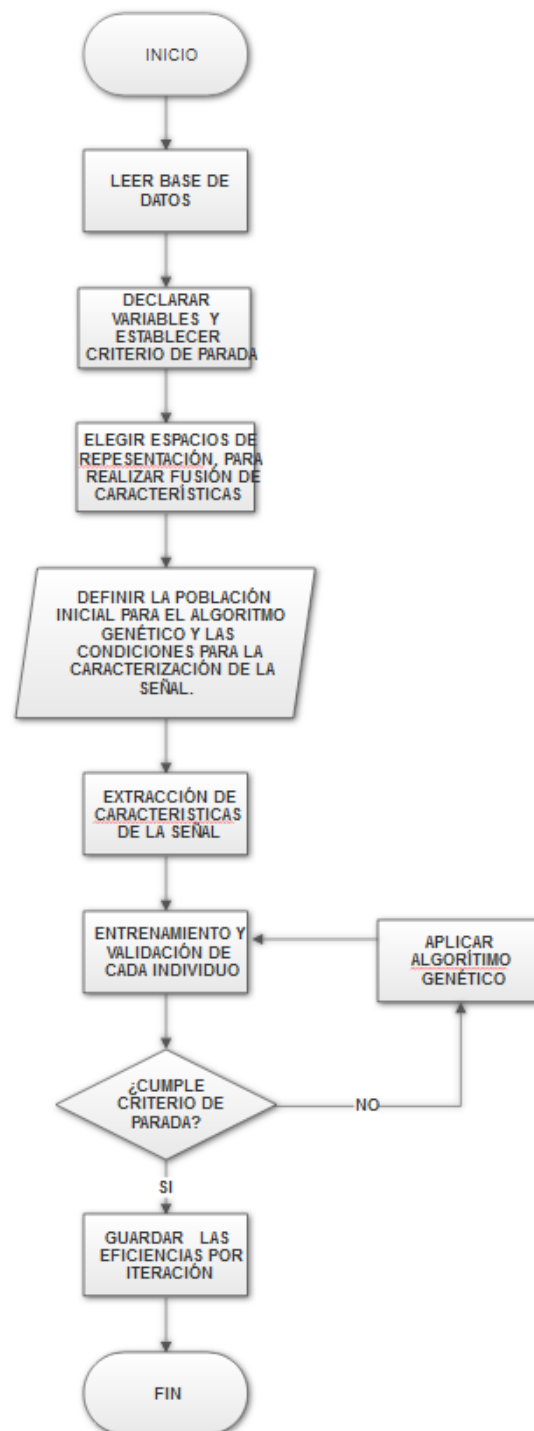


ILUSTRACIÓN 13. DIAGRAMA DE FLUJO ALGORITMO COMPLETO.

CAPÍTULO 14. RESULTADOS Y DISCUSIÓN.

14.1 Prueba I (Prueba algoritmo genético)

La prueba se efectuó con el fin de verificar el funcionamiento del algoritmo genético implementado, para esto se creó una población inicial de veinte individuos creados de manera aleatoria , a cada uno se le otorgó un peso (valor) diferente dado que aún no se había combinado con el algoritmo de validación, en base a los pesos el algoritmo estaba en la capacidad de identificar los individuos con mejores pesos (padres) y de esta manera extraer los mejores valores de la población (hijos), para de esta manera combinar los hijos y crear una nueva población.

14.2. Prueba II. (Prueba de libro de código).

La prueba se efectuó con el objetivo de encontrar el número óptimo de los K centroides de las señales de vibración de entrenamiento, con las cuales generó el libro de códigos. La prueba partió con la configuración inicial del algoritmo, particionado el conjunto total de vectores de características en 60% para entrenamiento y el 40% restante para validación además, fijando el número de estados en 2.

14.3. Prueba III (Análisis de tiempo de ejecución).

Teniendo en cuenta que las principales características que se buscan de un algoritmo son que pueda generar un resultado óptimo en un tiempo limitado, que este estructurado, que sea fácil de usar y sobre todo que sea eficiente (recursos y tiempo de computo que se requiere en su ejecución); resulta obvia la necesidad de generar algoritmos que permitan mantener tan bajo como sea posible el consumo de recursos del sistema. Se realizaron diferentes pruebas para analizar el tiempo de ejecución. En esta prueba se encontró que tanto el tiempo como la efectividad del algoritmo está directamente ligada a la cantidad de veces que se entrene y valide cada individuo.

14.4. Prueba IV. (Algoritmo completo)

La prueba IV se realizó con el fin de mostrar la eficiencia de los algoritmos genéticos a la hora de encontrar las mejores características en tiempo, frecuencia y tiempo-frecuencia, para la evaluación de la base de la base de datos de vibraciones del Bearing Data Center de la Western Reserve [32]. Para esta prueba se establecieron los parámetros que se muestran en la *tabla 6*.

Número de estados	2
Número de centroides	24
Número de Coeficientes	12
Número de filtros	24
Población	20
Porcentaje de mutación	5%
Número de veces que se entrenó y valido cada individuo (k)	25
Numero de iteraciones	100
Criterio de parada	La eficiencia de algoritmo no mejora en 4 iteraciones consecutivas

TABLA 6. PARÁMETROS DE LA PRUEBA IV.

Una vez establecidos los parámetros del algoritmo se procedió a ejecutar el programa, en cada una de las iteraciones se extrajo la mejor combinación de parámetros de tiempo, frecuencia, y MFCC. Además de la eficiencia media por iteración y la desviación estándar tal como se muestra en la *tabla 7*.

Ite rac ión	Características Tiempo	Características Frecuencia	Características Tiempo- Frecuencia	Eficiencia Media	Desviación estándar
1	[1 6 8 12 16]	[1 3 5 9 10 11 13 14 15 16]	[1 3 6 9 10 12]	0,9733333	0,03977532813915
2	[1 6 8 12 16]	[3 6 8 9 10 12 14 15 16]	[3 4 10 11]	0,9766666	0,03914883257893
3	[2 3 4 8 10 11 13 17]	[3 5 9 11 13 14 15]	[2 10 12]	0,9833333	0,02706115612056
4	[1 4 6 8 12 15 16]	[3 6 8 9 10 12 15 16]	[3 4 8 10 11]	0,9800000	0,03544934857793
5	[2 4 5 8 10 11 13 17]	[3 8 9 10 12 15 16]	[3 4 8 10 11]	0,9833333	0,02449489742783
6	[2 4 5 8 10 11 13 17]	[3 4 6 8 9 10 14 16]	[3 4 8 11]	0,9600000	0,03698584279832
7	[2 4 5 8 10 11 13 17]	[3 4 6 8 9 10 14 16]	[3 4 11]	0,9800000	0,03544934857793
8	[2 4 5 6 8 10 11 13 17]	[2 3 5 9 13 14]	[3 4 7 8 9 10 11]	0,97333333	0,03977532813915
9	[2 4 5 6 8 10 11 13 17]	[3 4 8 9 10 14 16]	[3 4 8 11]	0,97666666	0,02449489742783
10	[2 4 5 6 8 10 11 13]	[2 3 5 6 10 13 14]	[3 4 8 11]	0,98666666	0,01911941695938
11	[2 4 5 6 8 10 11 13 16]	[2 3 5 6 10 13 14]	[3 4 11]	0,97666666	0,04339108773912
12	[2 4 5 6 8 10 11 13 16]	[1 2 3 4 5 6 10 13 14]	[3 4 11 12]	0,98000000	0,02944737254926
13	[5 6 9 10 11 12 13 17]	[1 2 3 4 5 6 7 10 13 14]	[3 4 11 12]	0,990000000	0,02449489742783
14	[1 2 5 6 8 10 11 13]	[3 4 7 10 11 14 15 16]	[3 7 11]	0,97333333	0,035226121929
15	[5 6 9 10 11 12 13 17]	[3 4 7 10 11 14 15 16]	[3 4 11 12]	0,97333333	0,048989794855
16	[3 5 6 9 10 11 12 13 17]	[3 9 10 14 15 16]	[3 4 11 12]	0,98333333	0,034822853017
17	[1 2 4 5 6 8 10 11 13]	[3 9 14 15 16]	[3 6 7 8 11]	0,97333333	0,038015607270
18	[2 3 4 6 7 8 10 11 13 17]	[2 3 5 6 7 9 10 14]	[6 7 11]	0,98666666	0,026657887208
19	[1 2 6 8 10 11 13 15 17]	[5 9 10 14 15]	[3 4 7 11 12]	0,98333333	0,033063132149
20	[1 6 8 10 11 12 13 16 17]	[2 7 9 10 12 14 15]	[6 7 11]	0,99000000	0,024494897427
21	[1 2 6 8 10 11 12 16 17]	[2 3 5 6 7 9 14]	[6 7 11]	0,98333333	0,034822853017
22	[1 6 7 8 9 10 12 13 17]	[2 4 10 12 14 15]	[6 7 11]	0,98666666	0,020655911179

23	[1 2 3 4 5 6 7 8 10 12 17]	[2 3 5 6 7 8 9 14 16]	[2 3 4 7 11 12]	0,97666666	0,045150808607
24	[1 2 4 8 10 12 16 17]	[2 3 5 6 7 8 9 16]	[3 7 12]	0,97666666	0,034822853017
25	[1 6 7 9 10 13 16 17]	[2 3 5 6 8 9 12 14 15 16]	[2 3 4 9 11 12]	0,97333333	0,039775328139
26	[1 2 4 5 6 8 10 11 12 17]	[2 3 5 6 8 9 12 14 15 16]	[1 4 7 11 12]	0,98000000	0,042987818826
27	[1 2 4 5 6 8 10 11 12 17]	[2 3 5 6 9 12 14 15 16]	[1 4 7 11 12]	0,97333333	0,042987818826
28	[1 2 4 6 8 11 12 16 17]	[2 3 5 6 7 9 12 14 15 16]	[1 4 7 12]	0,97666666	0,037612338358
29	[1 2 4 6 8 10 11 12 16 17]	[2 3 5 6 9 12 14 15 16]	[1 4 7 11 12]	0,98000000	0,032659863237
30	[1 2 3 4 6 8 14]	[2 3 5 6 7 9 14 15 16]	[1 4 7 11]	0,96000000	0,042987818826
31	[1 2 3 4 5 6 8 10 11 12]	[1 2 4 5 6 9 11 12 15 16]	[1 4 7 11]	0,96000000	0,034748168525
32	[1 2 4 6 8 10 12 13]	[2 3 5 6 9 12 14 15]	[1 2 4 7 11]	0,98333333	0,024494897427
33	[1 2 4 5 6 8 10 11 12]	[2 3 5 7 8 9 15 16]	[1 4 7 8 9 11]	0,98000000	0,035449348577
34	[1 2 4 5 6 8 10 11 12 13 17]	[2 3 5 7 8 10 14 15 16]	[1 4 7 9 11]	0,97000000	0,042714124231
35	[1 2 4 5 6 8 9 10 11 12 13 17]	[2 5 6 7 9 15 16]	[3 4 7 8 9 11]	0,96333333	0,051152784636
36	[1 2 4 6 8 10 11 12]	[2 3 4 5 7 8 10 14 15 16]	[1 4 7 11 12]	0,98333333	0,040824829046
37	[1 2 4 5 6 9 11 12 13 17]	[2 3 4 5 7 8 10 14 15 16]	[1 4 7 9 10]	0,97000000	0,038783258564
38	[1 2 4 5 9 12 13 17]	[2 3 4 5 7 8 14 15 16]	[3 4 7 9 11 12]	0,97333333	0,032659863237
39	[1 2 4 8 9 10 11 12]	[5 6 9 12 13 15 16]	[1 4 7 9 10 12]	0,98000000	0,041228097958
40	[1 2 4 5 9 10 11 12]	[1 5 6 9 12 13 15 16]	[1 7 8 9 10]	0,99333333	0,016329931618
41	[1 2 3 4 5 8 9 10 11 12]	[1 3 5 6 9 12 13 15 16]	[3 4 7 9 10 12]	0,99333333	0,016329931618
42	[1 2 4 5 9 10 11 12]	[1 5 6 9 12 13 15 16]	[6 7 9 10 11]	0,99333333	0,016329931618
43	[1 2 4 5 6 9 10 11 12]	[1 5 6 9 12 13 15 16]	[3 4 7 9 10 12]	0,99333333	0,016329931618

TABLA 7. MEJORES VALORES POR ITERACION DE LA PRUEBA IV.

En la **Ilustración 10**. Se muestra el gráfico que representa la eficiencia por iteración en función de la iteración realizada.

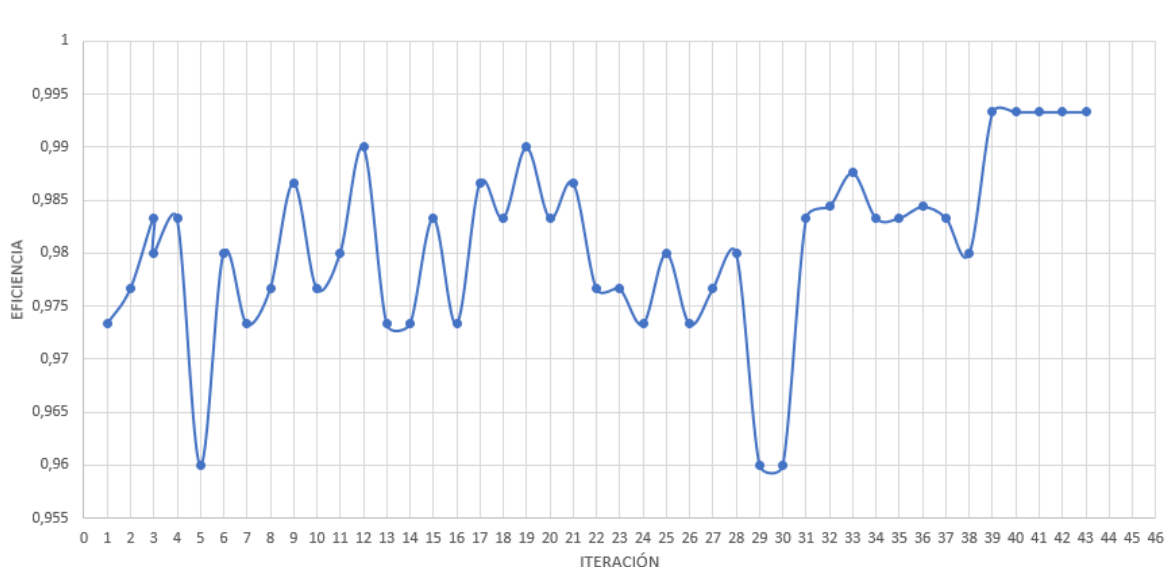


ILUSTRACIÓN 14 EFICIENCIA VS ITERACIÓN PRUEBA IV

Como se observa en la **Tabla 7**, la eficiencia del algoritmo para determinar qué características representan mejor el modelo, están fuertemente influenciadas por los espacios de representación tiempo y frecuencia. Sin embargo, los parámetros tiempo-frecuencia no aparentan influir de manera significativa, esto se puede apreciar en las iteraciones 40-41-42 y 43. En las cuales se presenta una tendencia con respecto a los dos primeros espacios de representación. Siendo las características de tiempo: **1. Media**, **2. Desviación estándar**, **4. Kurtosis**, **5. Pico Maximo**, **9. T4**, **10. Clearance Factor**, **11. Impulse Factor**, **12. Histogram upper bound**, y de frecuencia: **1. Media**, **5. Frecuencia Media**, **6. Desviación estándar**, **9. Inversa de la raíz cuadrática de la función por su media cuadrática**, **12. F14**, **13. F15**, **15. Desviación Media Absoluta** y **16. Rango inter cuartil**. Aquellas características que se repiten en las iteraciones que obtuvieron las mejores eficiencias. Por otra parte, las características tiempo-frecuencia no aparentan tener una tendencia definida ya que estas cambian notoriamente en las iteraciones mencionadas y la eficiencia no presenta ninguna alteración, aun así existen 3 características presentes en las mejores iteraciones las cuales son TF7, TF9 y TF10.

14.5. Prueba V. (Algoritmo completo).

A fin de corroborar los resultados obtenidos en la prueba IV se realizaron dos pruebas más, en cada una de ellas se cambiaron los parámetros de validación, específicamente en la cantidad de veces que se valida y entrena cada individuo y se disminuyó el porcentaje de mutación para observar el comportamiento de los datos y evaluar el desempeño del algoritmo genético, tal como se muestra en las *tablas 8 y 9*.

Número de estados	2
Número de centroides	24
Número de Coeficientes	12
Número de filtros	24
Población	20
Porcentaje de mutación	1%
Número de veces que se entrenó y valido cada individuo (k)	5
Numero de iteraciones	100
Criterio de parada	La eficiencia de algoritmo no mejora en 4 iteraciones consecutivas

TABLA 8. PARÁMETROS DE LA PRUEBA V

Una vez establecidas las condiciones iniciales del programa se procedió a realizar una segunda prueba del algoritmo completo, con el fin de comparar los resultados con la prueba anterior.

Iteración	Características Tiempo	Características Frecuencia	Características Tiempo-Frecuencia	Eficiencia Media	Desviación estándar
1	[1 2 3 5 7 9 11 12]	[1 5 6 9 12 13 14 15 16]	[1 2 3 6 12]	0,96666666	0,04680706864016
2	[1 8 10 12 13]	[4 5 6 7 9 10 12 16]	[1 2 3 6 12]	0,96000000	0,04731379838821
3	[1 2 5 7 9 10 11 12]	[4 5 6 7 9 10 12 16]	[1 4 5 6 7 8 9]	0,97000000	0,03482285301771
4	[1 2 8 10 12 13]	[4 5 6 7 9 10 12 16]	[1 2 3 6 12]	0,97000000	0,04361431438721
5	[1 2 5 7 9 10 11 12]	[4 5 6 7 9 10 12 16]	[1 4 5 6 7 8 9]	0,98000000	0,02065591117977
6	[1 2 5 7 9 10 11 12]	[4 5 6 7 9 10 12 16]	[1 2 3 6 12]	0,95333333	0,04298781882699
7	[1 2 5 7 9 10 11 12]	[4 5 6 7 9 10 12 16]	[1 4 5 6 7 8 9]	0,98666666	0,01911941695938
8	[2 5 7 9 10 11 12]	[4 5 6 7 9 10 12 16]	[1 5 6 7 8 9]	0,97666666	0,02882087698905
9	[1 2 5 7 9 10 11 12]	[4 5 6 7 9 10 12 16]	[1 5 6 7 8 9]	0,97333333	0,05177928019648
10	[1 2 3 5 7 9 10 11 12]	[4 5 6 7 9 10 12 16]	[1 4 5 6 7 8 9]	0,98000000	0,03544934857793
11	[2 5 7 9 10 11 12]	[4 5 6 7 9 10 12 16]	[1 5 6 7 8 9]	0,98000000	0,03265986323710
12	[1 2 3 5 7 9 10 11 12]	[2 4 5 6 7 9 10 11 12 16]	[1 5 6 7 8 9]	0,98000000	0,03265986323710
13	[2 5 7 9 10 11 12]	[4 5 6 7 9 10 12 16]	[1 5 6 7 8 9]	0,97666666	0,04361431438721
14	[1 3 5 7 9 10 11 12]	[4 5 6 7 9 10 12 16]	[1 5 6 7 8 9]	0,97333333	0,973333333333
15	[2 5 7 9 10 11 12]	[4 5 6 7 9 10 12 16]	[1 5 6 7 8 9]	0,97000000	0,047715316291
16	[2 5 7 9 10 11 12]	[4 5 6 7 9 10 12 16]	[1 5 6 8 9]	0,95666666	0,053719043329
17	[1 3 5 7 9 10 11 12]	[4 5 6 7 9 10 12 16]	[1 5 6 8 9]	0,97666666	0,024494897427
18	[1 5 7 10 11 12]	[4 5 6 7 9 10 12 13 16]	[1 6 8 9]	0,97333333	0,044017583299
19	[1 5 7 10 11 12]	[4 5 6 7 9 10 12]	[1 5 6 7 8 9]	0,98333333	0,03482285301
20	[1 2 5 7 10 11 12]	[4 5 6 9 10 12 16]	[1 5 6 8 9]	0,97666666	0,037612338358
21	[1 5 7 10 11 12]	[4 5 6 7 9 10 12 16]	[1 6 8 9]	0,96666666	0,038436356943
22	[1 5 7 8 10 11 12]	[4 5 6 7 9 10 12 16]	[1 6 8 9]	0,97666666	0,03761233835

23	[1 5 7 10 11 12 13]	[4 5 6 9 10 12 16]	[1 6 8 9]	0,98000000	0,042987818826
24	[1 5 7 10 11 12 13]	[4 5 6 7 9 10 12 16]	[1 6 8 9]	0,96333333	0,047471963658
25	[1 5 7 10 11 12]	[4 5 6 7 9 10 12 16]	[1 6 8 9]	0,97666666	0,043614314387
26	[1 5 7 10 11 12]	[4 5 6 9 10 12 16]	[1 6 7 8 9]	0,96333333	0,053942269977
27	[1 5 7 10 11 12 13 14]	[4 5 6 9 10 12 16]	[1 2 6 8 9]	0,97333333	0,038238833918
28	[1 5 7 10 11 12 13]	[4 5 6 9 10 12 16]	[1 7 8 9]	0,97666666	0,039710835507
29	[1 5 7 10 12 13]	[4 5 6 9 10 12 16]	[4 5 6 9 10 12 16]	0,97666666	0,028820876989
30	[1 5 7 10 11 12 13]	[4 5 9 10 12 16]	[1 2 6 8 9]	0,97333333	0,045554077519
31	[1 5 7 10 11 12 13 14]	[4 5 6 9 10 12 16]	[1 6 7 8 9]	0,97333333	0,035449348577
32	[1 5 7 10 11 12 13]	[5 6 10 12 16]	[1 6 7 8 9]	0,97000000	0,051152784636
33	[1 5 7 10 11 12 13]	[4 5 6 9 10 12 16]	[1 6 7 8 9]	0,96666666	0,038015607270
34	[5 7 10 11 12 13]	[5 6 9 12 16]	[1 6 7 8 9]	0,97000000	0,037144008068
35	[5 7 10 11 12 13]	[5 6 10 12 14 16]	[1 6 7 8 9]	0,98666666	0,032659863237
36	[5 7 10 11 12 13]	[5 6 9 16]	[1 6 7 8 9]	0,98000000	0,027687651680
37	[5 10 11 12 13 16]	[6 10 12 14 16]	[1 6 7 8 9]	0,98666666	0,019119416959
38	[5 7 11 12 13]	[5 6 9 16]	[1 6 7 8 9]	0,96666666	0,039306997849
39	[5 7 11 12 13]	[5 6 9 16]	[1 6 7 8 9 10]	0,97666666	0,031385384672
40	[5 10 11 12 13]	[5 6 9 16]	[1 6 7 8 9 10]	0,98666666	0,026657887208
41	[5 10 11 12 13]	[5 6 9 16]	[1 6 7 8 9 10]	0,986666	0,02665788720
42	[5 10 11 12 13]	[5 6 9 16]	[1 6 7 8 9 10]	0,98666666	0,045150808607
43	[5 10 11 12 13]	[5 6 9 16]	[1 6 7 8 9]	0,98333333	0,034822853017
44	[5 10 11 12 13]	[5 6 9 14 16]	[1 6 7 8 9]	0,99000000	0,03517565398288
45	[5 10 11 12 13]	[5 6 9 14 16]	[1 6 7 8 9]	0,99000000	0,01095445115010

46	[5 10 11 12 13]	[1 5 6 9 14 16]	[1 6 7 8 9]	0,983333333	0,03482285301771
47	[5 10 11 12 13]	[1 5 6 9 14 16]	[1 6 7 8 9]	0,966666666	0,04441763824072
48	[5 10 11 12 13]	[1 5 6 9 14 16]	[6 7 9 10 12]	0,980000000	0,04298781882699
49	[2 5 10 11 12 13]	[1 5 6 9 13 14 16]	[1 6 7 8 9 10]	0,973333333	0,05155605354839
50	[5 10 11 12 13]	[1 5 6 7 9 13 14 16]	[1 6 7 8 9]	0,980000000	0,03544934857793
51	[2 4 5 10 11 12 13]	[1 5 6 9 14 16]	[1 7 8 9]	0,963333333	0,04550360957276
52	[5 10 11 12 13]	[5 6 9 12 13 16]	[1 7 8 9]	0,983333333	0,01849292139916
53	[2 4 5 10 11 12 13]	[5 6 9 12 13 16]	[1 7 8 9]	0,973333333	0,03823883391876
54	[2 4 5 10 11 12 13]	[1 5 6 9 14 16]	[1 7 8 9]	0,973333333	0,047940294
55	[2 4 5 10 11 12 13]	[1 5 6 9 14 16]	[1 7 8 9]	0,983333333	0,03482285302
56	[2 4 5 10 11 12 17]	[5 6 9 12 13 16]	[1 4 7 8 9]	0,973333333	0,04401758330
57	[2 4 5 11 12 13]	[1 5 6 9 14 16]	[1 7 9]	0,9766667	0,03482285302
58	[2 4 5 11 12 13]	[1 5 6 14 16]	[1 7 8 9]	0,973333333	0,05087909004
59	[2 4 5 11 12 13 17]	[1 5 6 9 14 16]	[1 7 8 9]	0,973333333	0,03265986324
60	[2 4 5 11 12 13]	[1 5 6 9 14 16]	[1 7 8 9 12]	0,980000000	0,03265986324
61	[2 4 5 11 12 13 17]	[1 6 9 14 16]	[1 7 8 9]	0,983333333	0,03482285302
62	[2 5 11 12 13]	[1 5 9 14 16]	[1 7 8 9 12]	0,973333333	0,03977532814
63	[2 4 5 12 13]	[1 6 9 14 16]	[1 7 8 9 11 12]	0,983333333	0,02882087699
64	[2 4 5 11 12 13]	[1 6 9 14 16]	[1 7 8 9 12]	0,9866667	0,02665788721
65	[5 11 12 13]	[1 6 14 16]	[1 7 8 9]	0,9700000	0,04082482905
66	[5 11 12 13]	[1 6 12 14 16]	[1 7 8 9 12]	0,9866667	0,03265986324
67	[5 11 12 13]	[1 6 7 14 16]	[1 7 8 9 12]	0,9766667	0,04515080861
68	[2 5 11 12 13]	[1 6 9 14 16]	[1 5 7 8 9 12]	0,983333333	0,03306313215

69	[5 11 12 13]	[1 6 9 12 14 16]	[1 7 8 9 12]	0,9733333	0,04898979486
70	[5 11 12 13]	[1 3 5 6 8 13 14 16]	[1 4 5 7 8 9 12]	0,9800000	0,04122809796
71	[5 11 12 13]	[1 5 6 11 14 16]	[1 5 7 8 9 12]	0,9733333	0,03801560727
72	[5 10 11 12 13]	[1 3 5 6 8 11 13 14 16]	[1 4 5 7 8 9 12]	0,9766667	0,04339108774
73	[5 11 12 13]	[1 3 5 6 11 14 16]	[1 4 5 7 8 9 12]	0,9800000	0,04898979486
74	[5 10 11 12 13]	[1 3 5 6 11 14 16]	[1 4 5 7 8 9 12]	0,9766667	0,03138538467
75	[4 5 11 12 13]	[1 3 5 6 11 14 16]	[1 4 5 7 8 9 12]	0,9800000	0,02322041886
76	[5 10 11 12 13]	[1 3 5 6 11 14 16]	[1 4 5 7 8 9 12]	0,9900000	0,01849292140
77	[5 11 12 13]	[1 3 5 6 11 14 16]	[1 4 5 7 8 9 12]	0,9766667	0,04163136687
78	[11 12 13]	[1 3 5 6 11 14 16]	[1 4 5 7 8 9 12]	0,9700000	0,05150558560
79	[11 12 13]	[1 3 5 6 11 14 16]	[1 4 5 7 8 9 12]	0,9700000	0,05115278464
80	[5 11 12 13]	[1 3 5 6 11 14 16]	[1 4 5 7 8 9 12]	0,9900000	0,01849292140
81	[5 11 12 13]	[1 3 5 6 11 14 16]	[1 4 5 7 8 9 12]	0,9766667	0,02538340864
82	[5 11 12 13]	[1 3 5 6 11 14 16]	[1 4 5 7 8 9 12]	0,9800000	0,02768765168
83	[5 11 12 13]	[1 3 5 6 11 14 16]	[1 4 5 7 8 9 12]	0,9833333	0,03482285302
84	[5 11 12 13]	[1 3 5 6 11 14 16]	[1 4 5 7 8 12]	0,9800000	0,03544934858
85	[5 9 11 12 13 16]	[3 5 11 14 15 16]	[1 4 5 7 8 9 12]	0,9833333	0,03306313215
86	[5 9 11 12 13 16]	[3 5 11 14 15 16]	[1 3 4 5 7 8 12]	0,9700000	0,04171334026
87	[5 9 11 12 13 16]	[3 5 11 14 15 16]	[1 4 5 7 8 12]	0,9766667	0,03585261749
88	[5 9 11 12 13 16]	[3 5 11 14 15 16]	[1 4 5 7 8 10 12]	0,9700000	0,03161036233
89	[5 9 11 12 13 16]	[3 5 11 14 15 16]	[1 4 5 7 8 10 12]	0,9700000	0,04082482905
90	[4 5 9 10 11 12 13 16]	[3 5 11 14 15 16]	[1 4 5 7 8 9 12]	0,9766667	0,03761233836
91	[4 5 9 10 11 12 13 16]	[3 5 11 14 15 16]	[1 3 4 5 7 8 9 12]	0,9833333	0,02728438277

92	[5 9 11 12 13 16]	[3 5 8 11 14 15]	[1 4 5 7 8 9 12]	0,9733333	0,02944737255
93	[5 9 11 12 13 16]	[5 8 11 14 15 16]	[1 4 5 7 8 9 12]	0,9700000	0,03738911171
94	[5 9 11 12 13 16]	[5 8 11 14 15 16]	[1 4 5 7 8 9 12]	0,9733333	0,04298781883
95	[4 5 9 10 11 12 13 16]	[3 5 8 11 13 15 16]	[1 4 5 7 8 9 12]	0,9733333	0,02944737255
96	[5 9 11 12 13 16]	[3 5 6 11 14 15 16]	[1 4 5 7 9]	0,9866667	0,03265986324
97	[2 5 9 11 12 13 16]	[3 5 8 11 13 15 16]	[1 4 5 7 9]	0,9733333	0,04577730417
98	[2 5 9 11 12 13 16]	[3 5 8 11 13 15 16]	[1 4 5 6 7 8 9 10 12]	0,9766667	0,02449489743
99	[2 5 9 11 12 13 16]	[3 5 8 11 13 15 16]	[1 4 5 6 7 8 9 12]	0,9766667	0,04939306377
100	[2 5 9 11 12 13 16]	[3 5 8 11 13 15 16]	[1 4 5 6 7 8 9 10 12]	0,9766667	0,03738911171

TABLA 9. MEJORES VALORES POR ITERACIÓN PRUEBA V.

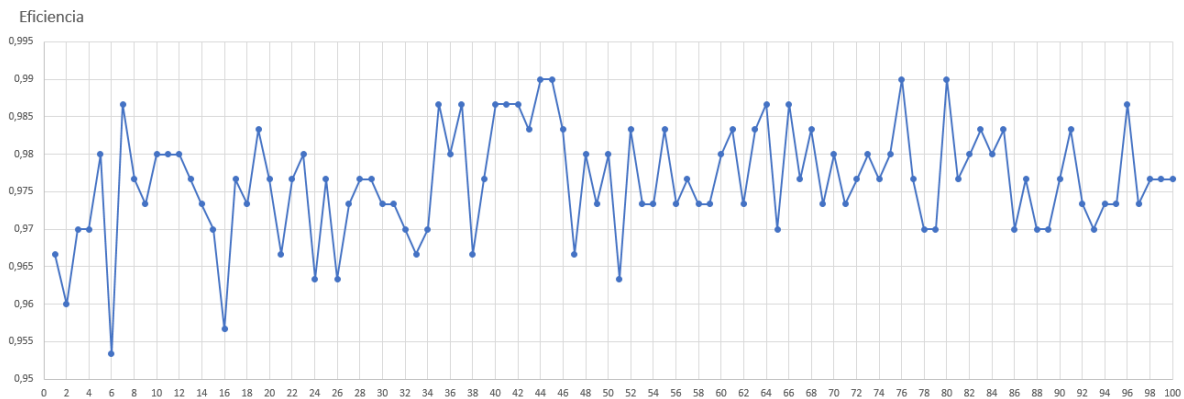


ILUSTRACIÓN 15. EFICIENCIA VS ITERACIÓN PRUEBA V

Al disminuir el número de veces que se entrenaba y validaba (k) cada individuo se esperaba una fuerte disminución en la eficiencia del algoritmo. Sin embargo, tal como se muestra en la **tabla 9**, a pesar de que la cantidad de veces que se entrenó y validó cada individuo (k) se redujo en una quinta parte, las eficiencias alcanzaron valores muy cercanos a los valores máximos obtenidos en la prueba IV. Aun así se aprecia cierta reducción con respecto al número de cifras significativas de la eficiencia. Esta reducción evidencio la fuerte relación que existe entre los tiempos de ejecución y el número de veces que se valide y entrene cada individuo (k) tal como se enunció en la prueba II, ya que en la prueba V el tiempo de ejecución fue 8 veces menor que en la prueba IV. De igual manera al reducir

el porcentaje de mutación se observa que el algoritmo tarda más iteraciones en volver a un valor optimo, esto se evidencia en las iteraciones 13-14-15-16-17 y 18, tal como se observa en la **Ilustración 12**. Debido al bajo porcentaje de cambio de características que existe entre una iteración y la siguiente, cuando el programa encuentra un óptimo local entrena y valida (k) varias veces características que aunque son óptimas locales no son buenas en comparación a los óptimos globales.

Tal como se muestra en la **tabla 9**, las mejores eficiencias se encuentran en las iteraciones 44-45-76 y 80. En las cuales se presenta cierta tendencia con respecto a las características que obtuvieron mejores resultados en los 3 diferentes espacios de representación siendo las características de tiempo **5. Pico** **Máximo 10. Clearance Factor 11. Impulse Factor 12. Histogram upper bound. 13. HISTOGRAM LOWER BOUND (HLB)**. De igual manera se observa que tanto en frecuencia como en tiempo frecuencia se sigue cierta tendencia al repetirse la gran mayoría de características en las iteraciones más significativas, siendo las características **5. Mediana 6. Desviación 14. Media geométrica y 16. Rango Inter cuartil** en la frecuencia y **TF1, TF7, TF8 y TF9** en tiempo-frecuencia.

14.6. Prueba VI. (Algoritmo completo)

En la prueba VI se sometió al algoritmo a las peores condiciones, con el fin de validar los resultados con las pruebas anteriores y observar los cambios que este podía presentar ante la modificación de los parámetros más significativos para el algoritmo genético tal como se muestra en la **tabla 10**.

Número de estados	2
Número de centroides	24
Número de Coeficientes	12
Número de filtros	24
Población	20
Porcentaje de mutación	1%

Número de veces que se entrenó y validado cada individuo (k)	1
Numero de iteraciones	100
Criterio de parada	La eficiencia de algoritmo no mejora en 4 iteraciones consecutivas

TABLA 10. PARAMETROS PRUEBA VI

Una vez establecidas las condiciones iniciales del programa se procedió a realizar una segunda prueba del algoritmo completo, con el fin de comparar los resultados con las pruebas anteriores.

Iteración	Características Tiempo	Características Frecuencia	Características Tiempo-Frecuencia	Eficiencia Media	Desviación estándar
1	[2 3 4 8 10 11 17]	[3 6 8 9 10 12 14 15 16]	[2 10 12]	0,96666667	0,038156763
2	[3 6 9 11 13 14]	[7 9 10 12 16]	[1 3 4 6 9 10 11 12]	0,96333333	0,055262058
3	[2 3 4 8 10 17]	[5 7 9 10 12 16]	[1 2 12]	0,97666667	0,051152785
4	[3 6 9 10 11 13 14]	[7 9 10 12 16]	[1 3 4 5 6 8 9 11 12]	0,95666667	0,046687514
5	[3 4 6 10 16 17]	[4 5 7 9 10 12 16]	[1 4 5 8]	0,96666667	0,056881066
6	[3 4 6 8 10 16 17]	[4 6 8 11 14 15]	[1 3 4 5 6 8 9 11 12]	0,96666667	0,033305022
7	[3 4 6 10 16 17]	[4 6 8 11 14 15]	[1 3 4 5 6 8 9 11 12]	0,95666667	0,052711397
8	[1 3 4 6 10 16 17]	[4 5 7 9 10 12 16]	[1 3 4 5 6 8 9 11 12]	0,96666667	0,05931775
9	[1 3 4 6 7 10 14 16 17]	[4 5 7 9 10 12 15 16]	[1 3 4 5 6 8 9 11 12]	0,96666667	0,04856679
10	[3 5 6 8 10 11 12]	[4 5 7 9 10 12 14 16]	[1 3 4 5 6 8 9 11 12]	0,98333333	0,034822853
11	[3 4 6 8 10 16 17]	[4 5 9 10 16]	[1 3 4 5 6 8 9 11]	0,96000000	0,053337892
12	[1 3 4 5 6 10 16 17]	[4 5 7 9 10 12 15]	[1 3 4 5 6 8 9 11 12]	0,96000000	0,062529736
13	[1 3 4 5 6 10 16 17]	[4 5 7 9 10 12 15]	[1 3 4 5 6 8 9 11 12]	0,97000000	0,059721019

14	[3 4 6 8 10 12 16 17]	[4 5 7 9 10 12 15]	[1 3 4 5 6 8 9 11 12]	0,96666667	0,051556054
15	[1 3 4 5 6 10 16 17]	[4 5 7 9 10 12 15]	[1 3 4 5 6 8 9 10 11 12]	0,96666667	0,044158739
16	[1 4 5 6 9 10 16 17]	[4 5 7 9 10 12 15]	[1 3 4 6 8 9 11 12]	0,97000000	0,046321729
17	[1 4 5 6 9 10 16 17]	[4 5 7 9 10 12 15]	[1 3 4 6 8 9 11 12]	0,97333333	0,036985843
18	[1 4 5 6 9 10 16 17]	[4 5 7 9 10 11 12 15]	[1 3 4 5 6 8 11 12]	0,96333333	0,054889964
19	[1 4 5 6 9 10 14 16 17]	[4 5 7 9 10 11 12 15]	[1 3 4 6 8 9 11 12]	0,97333333	0,030983867
20	[1 4 5 9 10 11 17]	[4 5 7 9 10 11 12 15]	[3 4 6 8 11 12]	0,97000000	0,05394227
21	[1 4 5 9 10 11 17]	[4 5 7 9 10 11 12 15]	[3 4 6 8 11 12]	0,97333333	0,038015607
22	[1 4 5 9 10 11 17]	[4 5 6 7 9 10 11 12 15]	[4 6 8 11 12]	0,98000000	0,035449349
23	[1 3 5 9 10 11 12 16]	[4 5 6 7 9 10 11 14 15]	[3 4 6 8 11 12]	0,98333333	0,033063132
24	[1 5 9 10 11 17]	[4 5 6 7 9 10 11 14 15]	[4 6 8 12]	0,96333333	0,05324109
25	[1 4 5 9 10 11 17]	[4 5 7 9 10 11 12 13 15]	[4 6 8 12]	0,96333333	0,039933493
26	[1 5 9 10 11 12 15]	[4 5 7 9 10 11 12 13 15]	[4 6 8]	0,98000000	0,036985843
27	[1 5 9 10 16 17]	[4 5 7 9 10 11 12 13 15]	[4 6 8]	0,95666667	0,047471964
28	[1 4 5 8 9 10 16 17]	[4 5 6 7 9 10 11 12]	[2 4 6 8 11 12]	0,96000000	0,049878306
29	[1 3 5 9 10 16 17]	[4 5 7 9 10 11 12 13 15]	[2 4 6 8 11 12]	0,95666667	0,06148074
30	[1 5 9 10 16 17]	[4 5 7 9 10 11 12 13 15]	[2 4 6 8 11 12]	0,96666667	0,045554078
31	[1 3 5 9 10 16 17]	[4 5 7 9 10 11 12 13 15]	[4 6 7 8 11 12]	0,95666667	0,062510505
32	[1 2 4 5 7 8 9 10 16 17]	[4 5 7 10 11 12 13 15]	[2 4 6 11 12]	0,96666667	0,040946248
33	[1 4 5 6 9 10 16 17]	[4 5 7 10 11 13 15]	[4 6 8 11 12]	0,97666667	0,043614314
34	[1 2 4 5 8 9 10 16 17]	[4 5 7 9 10 12 13 15]	[2 4 6 11 12]	0,96333333	0,049111214
35	[1 2 4 5 7 8 9 10 16 17]	[4 5 7 10 11 13 15]	[4 6 11 12]	0,97000000	0,037612338

36	[1 2 4 5 7 8 9 10 17]	[4 5 7 9 10 11 13]	[2 4 6 11 12]	0,97333333	0,056881066
37	[1 4 5 6 9 10 16 17]	[4 5 7 10 11 13]	[2 4 6 7 11 12]	0,97666667	0,031610362
38	[1 4 5 6 9 10 16 17]	[3 4 5 7 9 11 13]	[4 6 11 12]	0,96000000	0,04387633
39	[1 4 5 6 9 10 13 16 17]	[4 5 7 8 9 10 11 16]	[4 6 11 12]	0,96000000	0,04387633
40	[1 2 4 7 8 9 10]	[4 5 7 8 9 10 11 16]	[4 6 11 12]	0,98000000	0,032659863
41	[1 2 4 7 8 9 10]	[4 5 6 7 8 9 10 11 16]	[4 6 11 12]	0,95666667	0,044036815
42	[1 2 3 4 7 8 9 10]	[2 4 5 6 7 8 9 10 11 16]	[4 6 11 12]	0,96666667	0,047313798
43	[2 4 7 8 9 10]	[1 4 5 7 8 9 12 15]	[4 6 7 12]	0,97333333	0,039775328
44	[1 2 4 7 8 9 10]	[2 4 5 6 7 9 10 11 16]	[4 6 7 12]	0,97000000	0,042383441
45	[1 2 4 7 8 9 10]	[2 4 5 6 7 8 9 10 11 16]	[4 6 7 12]	0,96666667	0,05610526
46	[2 4 7 9 10 11 13]	[2 4 5 6 7 9 10 11 16]	[3 4 12]	0,97666667	0,051152785
47	[1 2 4 7 8 9 10 17]	[2 4 5 6 7 8 9 10 11]	[4 6 9 11 12]	0,96000000	0,053337892
48	[2 4 7 9 10 11 13]	[2 4 5 6 7 9 10 11 16]	[4 6 9 11 12]	0,96333333	0,040319753
49	[2 4 7 9 10 11 13]	[3 4 5 6 7 9 10 11 16]	[4 6 7 8 12]	0,97000000	0,047940294
50	[1 2 4 7 8 9 10 11 17]	[4 5 6 7 9 10 11 16]	[4 6 7 8 12]	0,95666667	0,064270226
51	[1 2 4 8 9]	[2 4 5 6 7 9 10 11 16]	[4 6 7 8 12]	0,96000000	0,064195541
52	[1 4 7 8 9]	[2 4 5 6 7 9 10 14]	[4 6 9 11 12]	0,97666667	0,051152785
53	[1 5 9 10 11 12 15]	[2 4 5 6 9 10 14]	[4 6 9 11 12]	0,98666667	0,032659863
54	[1 8 9 10 12]	[2 4 5 6 9 10 14]	[4 5 6 9 12]	0,95333333	0,048343563
55	[1 2 4 8 9 12]	[2 4 5 6 9 10 11 14]	[4 6 9 11 12]	0,97666667	0,043614314
56	[1 4 7 8 9]	[2 4 5 6 9 10 14]	[4 6 9 11 12]	0,97333333	0,042987819
57	[1 4 6 7 8 9 17]	[2 3 4 5 6 9 10 11 14]	[4 6 9 11 12]	0,97000000	0,051959322
58	[1 4 6 7 8 9 17]	[2 3 4 5 6 9 10 11 14]	[1 4 6 9 12]	0,95666667	0,051152785

59	[1 4 6 7 8 9 17]	[2 3 4 5 6 9 11 14]	[4 9 11 12]	0,97000000	0,047940294
60	[1 4 6 7 8 9 17]	[2 3 4 5 6 9 10 11 14]	[4 9 11 12]	0,96333333	0,045150809
61	[1 4 7 8 9 17]	[2 3 4 5 6 9 14]	[4 6 9 11 12]	0,97333333	0,045554078
62	[1 6 7 8 9 12 17]	[2 3 4 5 6 9 10 11 12 14]	[4 9 12]	0,97333333	0,041228098
63	[1 4 6 7 8 9 13]	[2 3 4 5 6 9 10 11 14]	[4 8 9 11 12]	0,97333333	0,051556054
64	[1 6 7 8 9 12 17]	[2 3 4 5 6 9 10 11 14]	[3 4 6 9 11 12]	0,94666667	0,048118585
65	[1 6 7 8 9 12]	[2 3 4 5 6 9 10 11 14]	[4 6 9 10 11 12]	0,97666667	0,037612338
66	[1 6 7 8 9 12]	[2 3 4 5 6 9 10 11 14]	[4 9 12]	0,96333333	0,053719043
67	[1 6 7 8 9 12 17]	[1 2 3 4 5 6 9 10 11 14]	[4 5 9 12]	0,96000000	0,05087909
68	[1 6 7 8 9]	[1 2 3 4 5 9 10 11 14]	[4 9 10 12]	0,96666667	0,045777304
69	[6 7 8 9 12 16]	[1 3 4 5 9 10 14]	[4 9 10 12]	0,98000000	0,035449349
70	[1 6 7 8 9]	[1 3 4 5 9 10 14]	[4 9 10 12]	0,97333333	0,039775328
71	[1 6 7 8 9 11]	[1 2 3 4 9 10 14]	[4 9 11 12]	0,95666667	0,056263425
72	[1 6 7 8 9]	[3 4 6 9 11]	[4 9 11 12]	0,96666667	0,036985843
73	[2 6 7 8 9 12]	[1 2 3 9 10 11 12 14]	[4 9 11 12]	0,97333333	0,04334062
74	[1 6 7 8 9 14 15]	[1 2 3 9 10 14]	[4 9 11 12]	0,96666667	0,066126264
75	[2 6 7 8 9 12]	[1 2 3 9 10 14]	[4 9 11 12]	0,96666667	0,041028319
76	[2 6 7 8 9 12]	[1 2 3 9 10 14]	[4 9 11 12]	0,97666667	0,043614314
77	[2 6 7 8 9 12]	[1 2 3 9 10 12 14]	[4 9 11]	0,97000000	0,059721019
78	[2 6 7 8 9 12]	[1 2 3 9 10 14 15]	[4 9 11 12]	0,97333333	0,065319726
79	[2 6 7 8 9 12]	[1 2 7 8 9 10 12 14]	[4 11 12]	0,97333333	0,051556054
80	[2 6 7 8 9 12]	[1 2 4 9 10 14 15]	[8 9 11 12]	0,96333333	0,052182549
81	[2 6 7 8 9 12]	[1 2 4 9 10 14 15]	[3 4 9 11 12]	0,97333333	0,041311822

82	[2 6 7 8 9 12]	[1 2 4 9 10 14 15]	[4 9 11 12]	0,96666667	0,047313798
83	[2 6 7 8 9 12]	[1 2 4 9 10 14 15]	[3 4 9 11 12]	0,97000000	0,044562008
84	[2 6 7 8 9 12]	[1 2 4 10 14 15]	[3 4 9 11 12]	0,97333333	0,04334062
85	[2 6 7 8 9 12]	[2 4 6 9 10 14 15]	[9 10 11 12]	0,97666667	0,042913134
86	[2 6 7 8 9 12 17]	[2 4 6 9 10 14 15]	[9 10 12]	0,96000000	0,04856679
87	[2 6 7 9 12 16 17]	[1 2 4 9 10 14 15]	[9 10 12]	0,97000000	0,043391088
88	[2 3 6 7 8 9 12]	[2 4 6 9 10 14 15]	[4 9 10 11 12]	0,97333333	0,05755803
89	[2 6 7 8 12]	[3 4 6 9 10 14 15]	[4 6 9 12]	0,95000000	0,055500882
90	[2 3 6 7 8 12]	[1 2 3 4 5 9 10 14 15]	[4 9 10 11 12]	0,96666667	0,048118585
91	[2 6 7 8 12]	[3 4 6 9 10 15]	[4 7 9 10 11 12]	0,96333333	0,053719043
92	[2 6 7 8 12]	[3 4 6 9 10 15]	[4 10 12]	0,97000000	0,053719043
93	[1 2 6 7 8 12 14 16]	[3 4 6 9 10 15]	[4 9 10 11 12]	0,97333333	0,039775328
94	[2 6 7 8 12 14 16]	[3 4 5 6 9 10 15]	[4 9 10 11 12]	0,96666667	0,042987819
95	[2 6 7 8 12 14 16]	[3 4 5 6 9 10 12 15]	[4 8 10 12]	0,97333333	0,038238834
96	[1 2 6 7 8 12 14 16]	[1 2 3 4 5 6 9 10 11 12 14 15]	[4 10 12]	0,97333333	0,05931775
97	[2 6 7 8 12 14 16]	[1 2 3 4 5 6 9 10 11 12 14 15]	[4 9 10 11]	0,97000000	0,047715316
98	[1 2 6 7 12]	[2 3 4 5 6 9 10 11 12 14 15]	[4 8 10 12]	0,97000000	0,049393064
99	[1 2 6 7 12 14 16]	[1 2 3 4 5 6 9 10 11 12 14 15]	[2 4 8 10 12]	0,96333333	0,055478764
100	[1 2 6 7 12 14 16]	[1 2 3 4 5 6 9 10 11 12 14 15]	[4 8 10 12]	0,96000000	0,041873825

TABLA 11.RESULTADOS PRUEBA VI

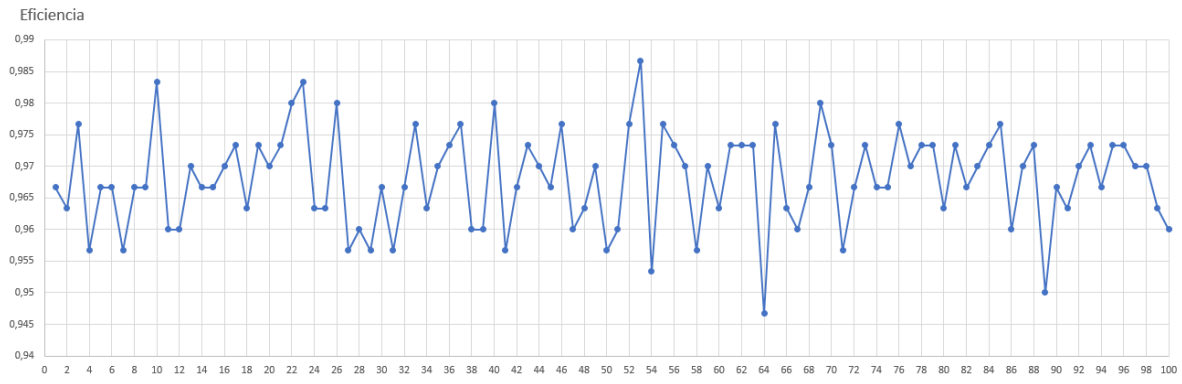


ILUSTRACIÓN 16.EFICIENCIA VS ITERACIÓN PRUEBA VI

Como era de esperarse al reducir la cantidad de veces que se entrena y valida cada individuo (k) se observa una reducción en la eficiencia de la función de validación, esto evidencia la importancia del ajuste de la función de validación (Fit value) para los algoritmos genéticos. Al mantener un porcentaje bajo de mutación se sigue corriendo el riesgo de que el algoritmo se quede en mínimos locales durante varias iteraciones consecutivas.

En la **tabla 12**. Se muestra una comparación de las características más representativas y con mejores resultados de cada una de las pruebas

Prueba	Características Tiempo	Características Frecuencia	Características Tiempo-Frecuencia	Eficiencia Media	Desviación estándar
Prueba IV	[1 2 3 4 5 8 9 10 11 12]	[1 3 5 6 9 12 13 15 16]	[3 4 7 9 10 12]	0,99333333	0,016329931618
	[1 2 4 5 9 10 11 12]	[1 5 6 9 12 13 15 16]	[6 7 9 10 11]	0,99333333	0,016329931618
	[1 2 4 5 6 9 10 11 12]	[1 5 6 9 12 13 15 16]	[3 4 7 9 10 12]	0,99333333	0,016329931618
	[1 2 3 4 5 8 9 10 11 12]	[1 3 5 6 9 12 13 15 16]	[3 4 7 9 10 12]	0,99333333	0,016329931618
Prueba V	[5 10 11 12 13]	[5 6 9 14 16]	[1 6 7 8 9]	0,99000000	0,03517565398288
	[5 10 11 12 13]	[5 6 9 14 16]	[1 6 7 8 9]	0,99000000	0,01095445115010
	[5 10 11 12 13]	[1 3 5 6 11 14 16]	[1 4 5 7 8 9 12]	0,99000000	0,01849292140
	[5 11 12 13]	[1 3 5 6 11 14 16]	[1 4 5 7 8 9 12]	0,99000000	0,01849292140

Prueba VI	[3 5 6 8 10 11 12]	[4 5 7 9 10 12 14 16]	[1 3 4 5 6 8 9 11 12]	0,98333333	0,034822853
	[1 4 5 9 10 11 12]	[4 5 6 7 9 10 11 12 15]	[4 6 8 11 12]	0,98000000	0,035449349
	[1 3 5 9 10 11 12 16]	[4 5 6 7 9 10 11 14 15]	[3 4 6 8 11 12]	0,98333333	0,033063132
	[1 5 9 10 11 12 15]	[4 5 7 9 10 11 12 13 15]	[4 6 8]	0,98000000	0,036985843
	[1 5 9 10 11 12 15]	[2 4 5 6 9 10 14]	[4 6 9 11 12]	0,98666667	0,032659863

ILUSTRACIÓN 17. COMPARACIÓN RESULTADOS

CAPÍTULO 16. CONCLUSIONES.

Al dar cumplimiento a los procedimientos propuestos con el fin de alcanzar los objetivos planteados en el capítulo 3. (Objetivos) y después de obtener los resultados expuestos en el capítulo 12 (Resultados y discusión) se puede concluir lo siguiente:

La etapa de parametrización y extracción de las características es una de las etapas más importantes a la hora de describir la señal de vibración, dado que al igual que con las técnicas de fusión se busca lograr una reducción en tiempos de computo, que permitan evaluar los estados de la maquina en la menor cantidad de tiempo, conservando la precisión y fiabilidad de la clasificación. Para esto se consideró desde un principio el uso de técnicas basadas en el cálculo de características en el dominio del tiempo, frecuencia y tiempo frecuencia de forma individual, para disminuir las pérdidas de información causadas por al parametrización de la señal.

Combinar las características extraídas de la señal estacionaria, por medio teorías de fusión permite establecer procedimientos óptimos para la detección de fallas y mejoramiento del proceso de extracción de características.

El uso de algoritmos genéticos en la extracción de características aumenta la fiabilidad del algoritmo de validación, ya que a partir de este se puede encontrar el conjunto de características más representativas del sistema, explorando casi por completo el espacio de solución, garantizando la convergencia del algoritmo en un valor óptimo global.

La eficiencia del algoritmo genético se encuentra altamente restringida a la fiabilidad de la función de validación que se use, dado que en cada iteración el algoritmo genético opera sobre los valores entregados por la función de validación tal como se muestra en el *capítulo 11*.

Implementar algoritmos genéticos y fusión de características en un mismo algoritmo garantiza la reducción de tiempos de computo de la mano con la exploración de todas las posibles soluciones incluyendo las soluciones no óptimas.

Los parámetros de validación y mutación son fundamentales para el uso de algoritmos genéticos ya que estos determinan los tiempo de ejecución, la eficiencia del sistema y la rapidez con la que se puede llegar a un valor óptimo. Y de esta manera encontrar un equilibrio entre eficiencia y tiempos de computo.

BIBLIOGRAFÍA

- [1]. Qin Hu; Aisong Qin; Qinghua Zhang; Guoxi Sun; Longqiu Shao, "Application of an information fusion method to compound fault diagnosis of rotating machinery," in Control and Decision Conference (CCDC), 2015 27th Chinese , vol., no., pp.3859-3864, 23-25 May 2015.
- [2]. Xiaofeng Liu; Lin Ma; Mathew, J., "Machinery Fault Diagnosis Based on Feature Level Fuzzy Integral Data Fusion Techniques," in Industrial Informatics, 2006 IEEE International Conference on , vol., no., pp.857-862, 16-18 Aug. 2006.
- [3]. Bartkowiak, R. Zimroz, Dimensionality reduction via variables selection – Linear and nonlinear approaches with application to vibration-based condition monitoring of planetary gearbox, Applied Acoustics, Volume 77, March 2014, Pages 169-177, ISSN 0003-682X
- [4]. M. Gutierrez, Mantenimiento Industrial efectivo, Envigado, Antioquia: Coldi, 2012.
- [5]. BILLUR B, BIRSEL A., "Fractional Fourier transform preprocessing neural networks and its application to object recognition [J]", 2002, vol. 15, no.1, pp.131-140
- [6]. KAMINSKA B, ARABI K, BELL I, "Analog and mixed-signal benchmark circuits - First Release [C]", 1997, pp.183-190
- [7]. SINGH N, SINHA A., "Optical image encryption using fractional Fourier transform and chaos [J]", Optics and Laser in Engineering, 2008, vol. 46, no.2, pp.117-123
- [8]. KUTAY M A, OZAKTAS H M, ARIKAN O, "Optimal filtering in fractional Fourier domains [J]", IEEE Tran Signal Processing, 1997, vol. 45, no.5, pp.1129-1143
- [9]. SHIN S G, JIN S L, SHIN S Y, "Optical neural network using fractional Fourier transform log-likelihood and parallelism [J]", Optics Communications, 1998, pp.4-6, pp.218-222
- [10]. Luo Hui; Youren Wang; Lin Hua; Jiang Yuanyuan, "Information fusion feature preprocessor based on FRFT for analog circuits fault diagnosis," in Electronic Measurement

- & Instruments (ICEMI), 2011 10th International Conference on , vol.4, no., pp.132-137, 16-19 Aug. 2011
- [11]. Mechefske, C.K., Objective machinery fault diagnosis using fuzzy logic. *Mechanical Systems and Signal Processing*, 1998. 12(6): p. 855-862.
 - [12]. Begg, C.D., et al., Dynamics modelling for mechanical fault diagnostics and prognosis. *Maintenance and Reliability Conference*, 1999. MARCON99.
 - [13]. Lee, S.K. and P.R. White, Higher-order time-frequency analysis and its application to fault detection in rotating machinery. *Mechanical Systems and Signal Processing*, 1997. 11(4): p. 637-650.
 - [14]. M. Gutierrez, *Mantenimiento Industrial efectivo*, Envigado, Antioquia: Coldi, 2012.
 - [15]. Baoping Tang, Tao Song, Feng Li, Lei Deng, Fault diagnosis for a wind turbine transmission system based on manifold learning and Shannon wavelet support vector machine, *Renewable Energy*, Volume 62, February 2014, Pages 1-9, ISSN 0960-1481.
 - [16]. Martin E. Liggins, David L. Hall, and James Llinas, *Handbook of Multisensor Data Fusion: Theory and Practice*, Second Edition, CRC Press, 2008.
 - [17]. S. K. Ambat, S. Chatterjee, K. V. S. Hari. "Fusion of algorithms for compressed sensing". *IEEE Transactions on Signal Processing*, vol.61, no.14, pp.3699–3704, July. 2013.
 - [18]. J. Yu and M. S. Yang, "A Generalized Fuzzy Clustering Regularization Model With Optimality Tests and Model Complexity Analysis," in *IEEE Transactions on Fuzzy Systems*, vol. 15, no. 5, pp. 904-915, Oct. 2007.
 - [19]. A. Bushong, "Fuzzy Clustering of Baseball Statistics," *NAFIPS 2007 - 2007 Annual Meeting of the North American Fuzzy Information Processing Society*, San Diego, CA, 2007, pp.66-68. doi: 10.1109/NAFIPS.2007.383812.
 - [20]. Genaro Daza Santacoloma "Metodología de *reducción de dimensión* para sistemas de reconocimiento".
 - [21]. Henry Alexander Blandón Hurtado, Yamir Fernando Martínez Mendoza "Desarrollo de una metodología para la detección de fallas a partir del análisis de vibraciones en rodamientos usando características de tiempo y modelos ocultos de markov (HMM)".
 - [22]. Rosaler, Robert C. *Manual del Ingeniero de Planta*. Mac-Graw-Hill. Interamericana de Editores 2002.
 - [23]. M. Fernandez Cabanas, M. Garcia Melero, G. A. Orcajo, J. M. Cano Rodríguez y J. Solares Sariego, *Técnicas para el mantenimiento y diagnóstico de máquinas eléctricas rotativas*, Barcelona: Marcombo, 1998.
 - [24]. V. Meruane, *Vibraciones Mecánicas*, Apuntes para el curso ME4701, Santiago de Chile: Universidad de Chile, 2011.
 - [25]. L. Muda, M. Begam, I. Elamvazuthi. "Voice Recognition Algorithms using Mel Frequency, Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques". *Journal of computing*. Vol. 2, Issue 3. ISSN: 2151-9617. March, 2010.
 - [26]. M. Carbonell, *TEORÍA DE LAS VIBRACIONES MECÁNICAS*, Madrid: Universidad Politécnica de Madrid, 2004 .
 - [27]. Diego Fernando Medina, *evaluación de metodologías para el diagnóstico automatizado de la severidad de fallas, a partir de vibraciones mecánicas en rodamientos, usando características de tiempo y de frecuencia, con modelos ocultos de Márkov*. Risaralda: Universidad Tecnológica de Pereira 2013.
 - [28]. Gale, «Baron Jean Baptiste Joseph Fourier,» de *Biography, Encyclopedia of World*, 2 ed., vol. 6, Detroit, Gale, 2007, pp. 32-33.
 - [29]. sinais, «Transformada de Fourier,» sinais, 2013. [En línea]. Available: http://www.sinais.es/Recursos/Curso-vibraciones/fundamentos/transformada_fourier.html. [Último acceso: 15 enero 2018].

- [30]. J. F. Ales Belsak, Adaptive Wavelet Transform Method to Identify Cracks in Gears, 2000 : Laboratory for Computer Aided Engineering, University of Maribor, Eslovenia.
- [31]. I. B. Moya, ANÁLISIS TIEMPO-FRECUENCIA DE LA SEÑAL DE VIBRACIÓN DE UN CAMBIADOR DE TOMAS EN CARGA, Leganés: UNIVERSIDAD CARLOS III DE MADRID,, 2009.
- [32]. C. W. R. University, «Bearing Data Center de la Western Reserve University,» Case School of Engineering , [En línea]. Available: <http://csegroups.case.edu/bearingdatacenter/pages/welcome-case-western-reserve-university-bearing-data-center-website>. [Último acceso: 16 enero 2018].
- [33]. O. C. Morales, Análisis tiempo-frecuencia de señales de vibraciones mecánicas para la detección de fallos en máquinas rotativas, Manizales: Universidad Nacional de Colombia, 2011.
- [34]. I. B. Moya, ANÁLISIS TIEMPO-FRECUENCIA DE LA SEÑAL DE VIBRACIÓN DE UN CAMBIADOR DE TOMAS EN CARGA, Leganés: UNIVERSIDAD CARLOS III DE MADRID,, 2009.
- [35]. G. Mosquera, Vibraciones mecánicas y su aplicación al mantenimiento predictivo, Caracas: Universidad central de Venezuela, 2000.
- [36]. A. PRO, «Detectando fallas en rodamientos utilizando métodos de Prueba eléctricos y mecánicos de vibración,» ALL TEST Pro, [En línea]. Available: www.sistemamid.com/download.php?a=79828. [Último acceso: 16 enero 2018].
- [37]. Lisa Sullivan, Wayne W. LaMorte, “Interquartile Range” Boston University School of Public Health [En línea]. Available: http://sphweb.bumc.bu.edu/otlt/mph-modules/bs/bs704_summarizingdata/bs704_summarizingdata7.html#. [Último acceso: 16 enero 2018]
- [38]. S. W. M. Yadav, «Vibration analysis of bearing for fault detection using time domain features and neural network.,» *International Journal of Applied Research in Mechanical* , vol. 1, p. 1, 2011..
- [39]. D. M. a. L. H, «feature selection for classification” Intelligent Data Analysis,» *Intelligent Data Analysis*, vol. 1, pp. 131-156, 1997.
- [40]. N. P. M. a. F. K, «A branch and bound algorithm for feature selecti on,» *IEEE* , vol. 26, pp. 917- 922, 1977.
- [41]. A. H. a. D. T. G, Learning with many irrelevant features, Cambridge: MIT Press, 1992, pp. 547- 552.
- [42]. K. K. a. R. L. A, The feature selection problem: Traditional methods and a new algorithm, Cambridge: MIT Press, 1992.
- [43]. C. Cardie, Using decision trees to improve case- based learning, Amherst: Morgan Kaufmann Publishers, 1993.
- [44]. L. H. a. S. R., «Feature selection and classification –A probabilistic wrapper approach,» de *Procedings of Ninth International Conference on Industrial and Engineering Applications of AI and ES*, Fukuoka, 1996.
- [45]. F. Martinez, G. Portale, H. Klein y O. Olmos, *Reconocimiento de voz, apuntes de cátedra para Introducción a la Inteligencia Artificial*.
- [46]. M. C. M. Joaquín Alegre Martín, INTRODUCCIÓN A LA ESTADÍSTICA DESCRIPTIVA PARA ECONOMISTAS, Palma: Universidad de las Islas Baleares , 2002.
- [47]. R. E. P. Guerra, *Tecnicas Estadisticas de clasificacion, un ejemplo de analisis cluster*, Hidalgo, 2006.

- [48]. Y. Gimenez, *Clasificación no supervisada: El método de k-medias*, Buenos Aires: Universidad de Buenos Aires, 2010.
- [49]. C. Garcia Camronero y R. I. Paredes portador, *Algoritmos de Aprendizaje: KNN & KMEANS*, Madrid: Universidad Carlos III de Madrid..